2006

Molecular genetic analyses of oilseed Brassica germplasm: determination of life forms and germplasm management strategies by using microsatellite markers and FLOWERING LOCUS-C (FLC1 and FLC3) gene sequences

Von Mark V. Cruz
Iowa State University

Molecular genetic analyses of oilseed *Brassica* germplasm: Determination of life forms and germplasm management strategies by using microsatellite markers and *FLOWERING LOCUS-C* (*FLC1* and *FLC3*) gene sequences

By

Von Mark V. Cruz

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Plant Breeding

Program of Study Committee:
Candice A. C. Gardner, Co-major Professor
E. Charles Brummer, Co-major Professor
Kendall R. Lamkey
John D. Nason
Mark P. Widrlechner

Iowa State University

Ames, Iowa

2006

Graduate College

Iowa State University

This is to certify that the doctoral dissertation of

Von Mark V. Cruz

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Committee member

Signature was redacted for privacy.

Committee member

Signature was redacted for privacy.

Committee member

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

Co-major Professor

Signature was redacted for privacy.

For the Major Program

TABLE OF CONTENTS

ABSTRACT vi

CHAPTER 1. GENERAL INTRODUCTION 1

Introduction 1

Dissertation organization 4

Literature review 4

Crop taxonomy and genetics 4

Molecular markers 8

Flowering time 11

Flowering-time genes 16

CHAPTER 2. CHARACTERIZATION OF FLOWERING TIME AND SSR MARKER ANALYSIS OF SPRING AND WINTER TYPE *BRASSICA NAPUS* L. GERMPLASM 23

Abstract 23

Introduction 25

Methodology 27

Results 31

Discussion 35

Summary and Conclusions 37

Acknowledgements 38

References 38

CHAPTER 3. SURVEY OF SEQUENCE VARIATION IN THE *FLOWERING LOCUS C* GENES (*FLC1* AND *FLC3*) IN ANNUAL AND BIENNIAL ACCESSIONS OF *BRASSICA NAPUS* L. 52

Abstract 52

Introduction 53

Materials and Methods 55

Results 59

Discussion 62

Conclusions 65

Acknowledgements 65

References 66

CHAPTER 4. ANALYSIS OF BULKED AND REDUNDANT ACCESSIONS OF *BRASSICA* GERMPLASM USING ASSIGNMENT TESTS OF MICROSATELLITE MARKERS 79

Abstract 79

Introduction 81

Methodology 83

Results 85

Discussion 88

Conclusion 90

Acknowledgements 91

References 91

CHAPTER 5. MEASURING THE EFFECTIVENESS OF ISOLATION OF *BRASSICA NAPUS* L. ACCESSIONS DURING CAGED GERMPLASM REGENERATION 102

Abstract 102

Introduction 103

Materials and Methods 105

Results and Discussion 108

Conclusion 110

Acknowledgements 111

References 111

CHAPTER 6. GENERAL CONCLUSIONS 119

REFERENCES CITED 123

ACKNOWLEDGEMENTS 137

ABSTRACT

Molecular markers were utilized in conjunction with phenological and morphological information and statistical methods for diversity analyses to address plant genetic resource conservation issues at the USDA-ARS North Central Regional Plant Introduction Station (NCRPIS). The objectives of this study can be categorized into two sets. The first set involves exploratory surveys of genetic variation by using molecular markers and DNA sequences to: (1) characterize nucleotide variation in two copies of the *FLOWERING LOCUS C* gene (*FLC1* and *FLC* 3) in *Brassica napus* L.; and (2) determine if variation in microsatellite markers and in the nucleotide sequence of the *FLC* copies can be associated with life forms in *B. napus*. The second set of specific objectives addresses good genebank practices and includes: (1) characterization of flowering time of the spring-type *B. napus* without vernalization; (2) assessment of the validity of bulking germplasm by examining selected sets of *B. rapa* L. bulks; (3) determining whether bulking of putative duplicates or recently identified redundant accessions of *B. napus* is appropriate; and (4) determining if pollen can pass through the screen mesh used in regeneration of *B. napus* and cause cross-contamination between adjacent field cages. The first set of objectives was addressed by analyzing 33 microsatellites from 18 *B. napus* linkage groups and three expressed sequence tags (ESTs), and by examining short (exon 4 to 6) and long (exon 2 to 7) sequences of *FLC1* and *FLC3* in 50 representative accessions from the NCRPIS *B. napus* collection. It was determined that groups of life forms, as well as probable duplicate accessions, can be distinguished by using microsatellite markers. Eleven microsatellite loci displayed significant associations with flowering time. It was also determined that there are relatively weak associations between derived genetic distances and geographic origins and between derived genetic distances and life forms. Analysis of the *FLC* copies indicated that there is low polymorphism in exons 4 to 6 in both annual and biennial types of the crop. A life-form specific haplotype in either copy of *FLC* analyzed was not found. The curatorial concerns

addressed by the second set of objectives were investigated by utilizing a set of ten microsatellites selected from those above, and by bioassay of a dominant herbicide-resistance trait in putative progenies of transgenic and non-transgenic *B. napus*. To resolve issues in germplasm bulking in *B. rapa* and the problem of duplicate accessions in *B. napus*, information derived from the ten microsatellites was analyzed by using assignment tests together with tests for population differentiation and cluster analysis of genetic distances. Results from this study indicated the presence of unique accessions among *B. rapa* bulks and in the *B. napus* putative duplicates. It was demonstrated that use of genetic probability profiles of component accessions could corroborate cluster analyses while providing additional information. A dominant herbicide-resistance trait was useful in determining the presence of low-level cross-contamination among field cages during regeneration. The study confirmed that wind-blown *B. napus* pollen could pass through the current screen mesh size used during regeneration and cause cross-fertilization in adjacent plots, but at a very low and tolerable level.

CHAPTER 1. GENERAL INTRODUCTION

Introduction

A primary objective of the conservation of plant genetic resources is to make genes and genotypes of crops available to current and future plant improvement programs. Plant germplasm conservation aims to safeguard against the loss of genetic diversity, preserve genes that have potential in developing better varieties of plants, and ensure their availability to germplasm users. The widespread replacement of traditional plant cultivars due to the adoption of high yielding varieties during the last century coupled with the disappearance of wild crop relatives due to human activities, habitat destruction, and threats from climate change or because of natural disasters, have highlighted the importance of conserving plant genetic resources.

Many of the world's plant genetic resource centers, or genebanks, were established in the 1960s to preserve germplasm of crops and their weedy or wild relatives (Koo et al., 2004). However, the establishment of such collections began in the late 1800's by the Russian Bureau of Applied Botany (now the N.I. Vavilov Institute of Plant Industry or VIR) and the United States Department of Agriculture. These early collection activities were primarily aimed at providing materials to crop improvement programs and for direct introduction of new crops.

Scientific approaches to conservation were strengthened with the establishment of the International Board for Plant Genetic Resources (now the International Plant Genetic Resources Institute) (Ford-Lloyd, 2001). Standards were set and descriptors for major crops were formulated to enable systematic characterization and evaluation of germplasm. Morphological and phenological germplasm characterization is often supplemented with genotypic information derived from molecular markers based on protein and/or DNA polymorphisms (Spooner et al., 2005). Results from these diversity analyses have proven useful in gap analysis to help establish priorities for the

acquisition of germplasm, determine the need for further characterization, and validate regeneration methods (Ramanatha Rao and Hodgkin, 2002).

For germplasm users and curators, the availability of phenotypic and genotypic data is critical to the identification of specific germplasm of interest to meet research objectives. Genebank curators are likewise able to assess the amount of existing variation on hand, which is important in determining conservation needs and setting future priorities (Ramanatha Rao and Hodgkin, 2002). The studies presented in the following chapters utilize molecular-marker, phenological and morphological information in conjunction with statistical methods for diversity analyses to address plant genetic resource conservation issues. Specifically, uses of microsatellite markers and gene sequences to address curatorial management questions pertaining to the *Brassica napus* and *B. rapa* collections at the USDA-ARS North Central Plant Introduction Station (NCRPIS) were investigated.

Brassica napus and *B. rapa* are collectively known as 'rapeseed' in Australia, Canada and the United States, and as 'oilseed rape' in the United Kingdom. Both species are important crops that were originally grown as sources of lamp oils in the 13th century and lubricants during the industrial revolution (McNaughton, 1995). Their use as sources of edible oil is recent (Downey, 1964).

The importance of rapeseed as an oil crop is reflected by the continued increase in its worldwide production. In 2005, it was reported that 54 countries produced 46 million metric tons (Mmt) of rapeseed, up by 21% from 2003 (FAO, 2006). Currently, China is the greatest producer (13 Mmt) followed by Canada (8.5 Mmt), India (6.4 Mmt) and Germany (4.7 Mmt). The United States ranks 10th in rapeseed production; more than 91% of U.S. rapeseed hectarage is in Minnesota and North Dakota (USDA-NASS, 2006). This is probably due to favorable seasonal temperature, moisture and photoperiod conditions during the growing season and to close proximity to Canada's agricultural and rapeseed oil processing infrastructure.

In 2001, about 86,727 accessions of *Brassica* were conserved in 140 genebanks worldwide (IPGRI, 2001). The major repositories, having 5000 or more accessions, are the Institute of Crop

Germplasm Resources and the Institute of Oil Crops Research in China, the Haryana University in India, and the Institute for Plant Genetics and Crop Plant Research in Germany (IPGRI, 2001).

In the United States, the USDA-ARS NCRPIS conserves about 1300 accessions of rapeseed. There are 600 accessions of *B. napus* and about 700 accessions of *B. rapa* (USDA-GRIN, 2006). The research presented here was conducted to address curatorial concerns pertaining to the management of the rapeseed collection, and was designed specifically to address three questions about the rapeseed collections at NCRPIS: (1) can molecular markers be used to determine the life form or flowering time of *B. napus* accessions before regeneration in the field? (2) are the putative duplicate accessions of *B. napus* and component accessions of bulked *B. rapa* germplasm accessions genetically homogeneous? and (3) does the current method of *Brassica* regeneration using screen cages prevent pollen contamination between adjacent cages?

The specific objectives of this study can be categorized into two sets. One set pertains to exploratory surveys of genetic variation using molecular markers and DNA sequences to: (1) characterize nucleotide variation in two copies of the *FLOWERING LOCUS C* gene (*FLC1* and *FLC* 3) in *B. napus*; and (2) determine if variation in microsatellite markers and at the nucleotide level in *FLC1* and *FLC* 3 can be associated with life forms in *B. napus*. The second set of specific objectives addresses good genebank practices and includes: (1) characterization of flowering time of the spring-type *B. napus* without vernalization; (2) assessment of the validity of bulking germplasm by examining selected sets of *B. rapa* bulks; (3) determining whether bulking of putative duplicates or recently identified redundant accessions of *B. napus* is appropriate; and (4) determining whether pollen can pass through the screen mesh used in regeneration of *B. napus* and cause cross-contamination between adjacent field cages.

Dissertation organization

This dissertation consists of this general introduction, four papers and a general conclusions section. The first paper describes the results of characterization of flowering time and the application of microsatellite markers to life-form determination. The second determines if natural variation that can be associated with flowering time exists in two copies of the *FLC* gene (a central gene in the autonomous pathway of flowering). The third paper deals with issues surrounding the practice of bulking *Brassica* germplasm and introduces an analytical method used in conservation biology to determine differences among putative duplicates. The fourth paper examined if pollen flow occurs between caged accessions of *B. napus*; screen cages are used to prevent contamination of accessions during the regeneration process. Molecular markers were not used in the fourth paper, rather, a simple bioassay of the progeny using an herbicide-resistant trait was conducted. Each paper was written in a format following the guidelines for each journal where it was or will be submitted. The papers are followed by a General Conclusions section and full citations to the references cited in the general introduction and conclusions.

Literature review

Crop taxonomy and genetics

Brassica napus and *B. rapa* belong to the plant tribe Brassiceae, family Brassicaceae (formerly Cruciferae), Subclass Dilleniidae, Class Magnoliopsida (Dicotyledons). Members of the tribe Brassiceae can be distinguished by the presence of conduplicate cotyledons (cotyledons that are longitudinally folded around the radicle) and/or heteroarthrocarpic (transversely segmented) fruits (Warwick and Sauder, 2005). Some of the commonly known close relatives of *B. napus* and *B. rapa*

include the vegetables, cabbage, cauliflower, broccoli, brussels sprouts, kohlrabi, and kale within *B. oleracea* L.

There has been some confusion on the nomenclature of *B. rapa*. Much of the literature refers to two scientific names (*Brassica rapa* and *B. campestris*) depicting this species. Confusion resulted when *Brassica rapa* and *B. campestris* were described as separate species by Linnaeus. He noted that *B. rapa* is the turnip form and *B. campestris* the wild weedy form. In 1833, it was subsequently determined that these were the same species, and the two taxa were combined under *B. rapa* (Toxeopus et al., 1984). In current publications, rapeseed researchers still cite both names, with *B. rapa* still being commonly used to denote the turnip rape forms.

There are six major species of cultivated *Brassica*. The evolution of these cultivated species was inferred from results of chromosome counts and chemotaxonomy (Vaugh, 1977). Relationships among these species are frequently presented following the 'Triangle of U' (Figure 1), named after the Japanese scientist who postulated the interspecific relationships after evaluating interspecific crosses. The U triangle shows that hybridizations among the three different genomes of the diploid progenitor species - *B. rapa, B. nigra* L. (Koch), and *B. oleracea* – gave rise to the amphidiploid species – *B. juncea* (L.) Czern., *B. carinata* Braun, and *B. napus.* It was hypothesized that the amphidiploids arose spontaneously from overlap of cultivation areas of the diploid species. Studies have confirmed that it is possible to produce artificial (resynthesized) *B. napus* and *B. juncea* by doing reciprocal crosses between their respective parent species (Butruille et al., 1999; Tsunoda et al., 1980). However, progenies from the interspecific crosses of *Brassica* exhibit partial fertility or sterility because of different chromosome numbers of the cultivated species. Progeny resulting from pollinations typically fail to develop due to a lack of endosperm development (Plant Biosafety Office, 1999).

In rapeseed breeding, *B. napus* and *B. rapa* are known as Argentine and Polish types, respectively (Potts et al., 1999). *Brassica napus* is an amphidiploid (AACC genome, 2n=38) and is

believed to have arisen from interspecific hybridization between the diploid species *B. rapa* (AA genome, 2n=20) and *B. oleracea* (CC, 2n=18) (Gómez-Campo, 1999; Sauer, 1993). *Brassica rapa* has been presumed to be the oldest species among the cultivated *Brassica* and probably evolved in Europe. Forms of *B. rapa* can still be found in weedy habitats in areas extending from the western Mediterranean to Central Asia. This wide distribution accounts for the greater genetic variability found in *B. rapa* compared to *B. napus*. The latter species is presumed to have developed out of cultivation. No wild forms of *B. napus* have been found in nature, except in cases where it occurs as an escape (Gómez-Campo and Prakash, 1999). It is believed that the biennial type of *B. napus* is older than the annual form. The annual types were probably developed through selection for early flowering by humans (Lackey, 1996; Gómez-Campo and Prakash, 1999).

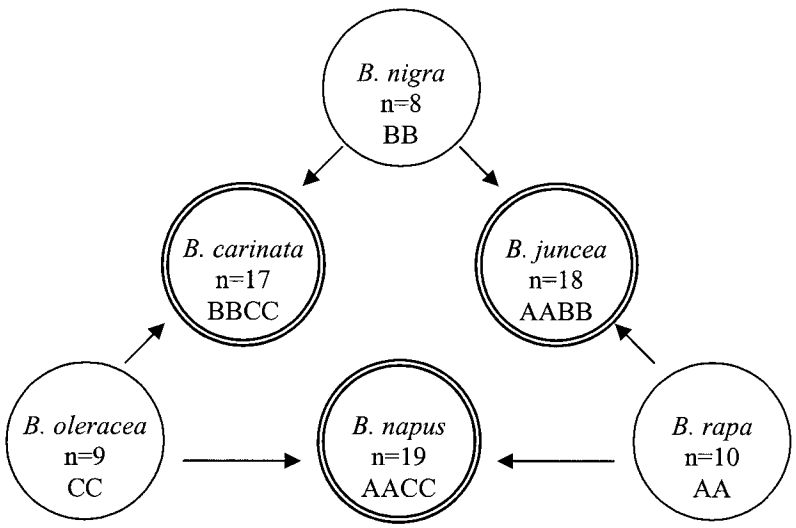


Figure 1. Genomic relationships among the *Brassica* species (redrawn from Kimber and McGregor, 1995).

Active public and commercial breeding programs for rapeseed are conducted in Canada, China and several institutions in Europe (Philips and Khachatourians, 2001). Breeding efforts aim to improve oil and meal qualities, yield, disease and herbicide resistance, and agronomic traits such as

winter hardiness and seed color (Rahman, 2001; Rife and Salgado, 1996; Rife and Zeinali, 2003; Tsunoda et al., 1980; Shahidi, 1990). The development of *B. napus* cultivars that are early maturing and higher yielding led to the replacement of the Polish type *B. rapa.* Varieties from the latter species however, have more shatter-resistant pods and are less likely to produce 'green seed', which is undesirable in oil production (Falk et al., 1998).

Uses of *B. napus* and *B. rapa* rapeseeds

Swedes, a form of *B. napus*, are commonly used as animal fodder in Europe (McNaughton, 1995). Oilseed *B. napus* crop varieties are primarily used in the production of oil for food and non-food applications (McNaughton, 1995). In the past, rapeseed oil was mainly for industrial use with limited edible applications. The development of low erucic acid varieties of rapeseed and turnip rape in the 1960s through traditional plant breeding led to production of edible oil and its widespread use (Downey, 1964; Steffansson et al., 1961). Erucic acid is an unsaturated fatty acid that cannot be digested by most animals and has been documented to accumulate in the human body and can damage the heart and adrenal organs (Guil et al., 1997). It occurs at about 22-66% in unselected rapeseed oil making it undesirable for human and animal consumption (Thomas, 2003). Currently, there are *B. rapa* lines that have very low to zero levels of erucic acid in the oil (Rahman, 2002). Such varieties with low erucic acid are also known as LEARs (for Low Erucic Acid Rapeseed).

In addition to having low erucic acid, many current rapeseed varieties were developed to have low glucosinolates, enabling the meal residue from the oil extraction process to be suitable for animal feed. The varieties with low levels of both erucic acid and glucosinolates are called canola to designate higher quality rapeseeds. 'Canola' was coined in 1978 as a Canadian trademark to distinguish rapeseed products with less than 2% erucic acid and less than 30 μmol g^{-1} meal of aliphatic glucosinolates, and are also referred to as 'double low' varieties. In the NCRPIS collection,

many accessions have canola quality traits. In a sample of 364 accessions of *B. napus* and *B. rapa* that was analyzed, about 25% have 0.0 to 9.9% erucic acid, and 40% have low glucosinolates (Auld et al., 1988).

Canola oil is primarily utilized as a healthier alternative to other vegetable oils because of its higher nutritional value as indicated by significant amounts of essential fatty acids (linoleic and linolenic) and Vitamin E (Hu et al., 1999; Singh et al., 2001). There are rapeseed varieties noted for their specific fatty-acid contents, called HOAR (for High Oleic Acid Rapeseed) and LLAR (Low Linolenic Acid Rapeseed) (Thomas, 2003).

For non-food applications, rapeseed types that are classified as 'high erucic acid (HEAR)' varieties may serve as good sources of industrial lubricants, bath and hair oils, and biodiesel for farm machinery and trucks (Kimber and McGregor, 1995; Peterson et al., 1999). HEAR varieties may also be used in the production of rubber and plastics (Kershaw, 1998). Other non-food products from oilseed *Brassica* include important protein concentrates and isolates that are utilized as emulsifiers, surfactants and hydrophobic biopolymers in industrial production (BioMat.Net, 2001). The production of oils with various properties for industrial purposes is being studied in *Brassica* by using biotechnology either through manipulation of the chain length or functionality of fatty acids, or the insertion of genes for novel fatty acids from other plant species (Murphy, 2004).

Molecular markers

Characterization with biochemical and molecular markers plays a key part in plant genetic resource management by providing supporting information for use in conjunction with morphological description to determine the level and nature of diversity of germplasm (Bretting and Widrlechner, 1995). A primary objective of the management of plant genetic resources is to conserve and utilize as much genetic diversity as possible, and maintain this diversity through time (Esquinas-Alcázar,

2005). Understanding the nature of a collection's diversity is also very important and of interest to crop improvement programs (Mohammadi and Prasanna, 2003).

Numerous molecular markers are used to guide germplasm management and answer questions in genetic conservation (Spooner et al., 2005). Current molecular markers can be grouped into three general types: protein variants, DNA sequence polymorphisms, and DNA repeat variation (Schlötterer, 2004). Each type of molecular marker has its own advantages and disadvantages, and the choice of what to use depends on the nature of the research question or problem (Farooq and Azam, 2002). In the past, seed oil fatty-acid profiles, level of glucosinolates, and isozyme proteins from cotyledons served well as molecular markers in the comparison of *Brassica* cultivars (Adams et al., 1989; Hackenberg and Kohler, 1996; Mundges et al., 1990; White and Law, 1991). Isozymes are allelic variants of enzymes coded by structural genes. Isozymes have been also used in estimating the outcrossing rate in *B. napus* (Becker et al., 1992) and in testing the effectiveness and validity of germplasm regeneration procedures (Diaz et al., 1997). Use of isozyme markers in *Brassica,* as well as in many other crops, are gradually being superseded by DNA-based markers, especially those derived from the polymerase chain reaction (PCR), enabling analyses from small amounts of DNA and allowing the study of greater amounts of variation.

Restriction fragment length polymorphism (RFLP) was the first DNA-based marker technique used to characterize *Brassica* germplasm (Charters et al., 1996). RFLP markers are bands that correspond to genomic DNA fragments, about 2-10 kb, resulting from digestion of linear DNA by endonucleases (Spooner et al., 2005). RFLPs are popular in QTL analyses in *Brassica* and have been used to map genes controlling vernalization requirements (Ferreira et al., 1995) and glucosinolate content (Uzunova et al., 1995). Other DNA-based markers that have been used in *Brassica* include randomly amplified polymorphic sequence (RAPD), amplified fragment length polymorphism (AFLP), cleavable amplified polymorphic sequence (CAPS), simple sequence repeat (SSR or microsatellite), and more recently, single strand conformational polymorphism (SSCP) (Pires

et al., 2004) and single nucleotide polymorphism (SNP) (Gupta et al., 2004). In general, recently developed DNA-based techniques, such as SSRs, DNA sequencing, and SNPs, offer high-throughput analysis of genotypes and have been increasingly utilized (Khlestkina and Salina, 2006; Schlötterer, 2004). The characteristics of the most commonly used markers, as well as their advantages and disadvantages are listed in Table 1.

Table 1. Advantages and disadvantages of most commonly used DNA markers (Avise, 2004; Collard et al., 2005; Schlötterer, 2004).

Marker	Advantage	Disadvantage
Isozymes	• Inexpensive • Universal protocols	• Limited number of loci assayed • Moderate resolution
RFLP	• Robust and reliable • Transferable across populations	• Long laboratory development time • Time consuming and laborious • Large amounts of DNA required • Limited polymorphism in related lines
RAPD	• Quick and inexpensive assay • Small amount of DNA required • Multiple loci from single primer possible	• Poor reproducibility • Problems in marker transferability
AFLP	• Multiple loci • High level of polymorphism	• Large amounts of DNA required • Complicated methodology
SSR	• Robust and reliable • Technically simple • Transferable across populations	• Usually require polyacrylamide electrophoresis
SNP	• High abundance • Easy to type/score • Cross study comparisons are easy	• Expensive to isolate • Ascertainment bias • Low information content of a single SNP

Microsatellites, also termed simple sequence repeats (SSR) and short tandem repeats (STR), correspond to regions in the DNA composed of repeating units of nucleotides between one and 10 base pairs in length (Ciofi et al., 1998). Variations observed in microsatellite loci are due to the differences in the number of repeat units caused by slippage events during DNA replication or by point mutations (Kruglyak et al., 1998). Microsatellites have gained more popularity than have other DNA markers because of several advantages, including the increasing availability of primer sequences for specific crops in public databases. In addition, many *Brassica* SSRs are able to cross amplify between the cultivated species (Lowe et al., 2002; Plieske and Struss, 2001).

Microsatellites most often behave in a co-dominant fashion, allowing detection of heterozygotes. This makes microsatellite markers very suitable for population genetics and phylogenetic studies requiring computation of gene frequencies. SSRs are becoming very useful in marker-assisted selection in plant breeding (Francia et al., 2005). SSR markers are developed through the laborious task of constructing and sequencing genomic libraries. Because *Brassica* EST and genomic sequence data are currently available in the public domain, the number of SSRs is expected to increase as the search for repeat motifs is conducted *in silico* (Love et al., 2004; Tonguç and Griffiths, 2004).

Flowering time

Determination of plant flowering time results from the interaction among genetic, environmental and physiological factors (Werner et al., 2005). From an evolutionary perspective, flowering should occur when the plant can maximize its chance of propagating (Casci, 2000). Because of the complexity of flowering mechanisms, such mechanisms have been the focus of many studies for several decades. Understanding these mechanisms offers the prospect of manipulating flowering time and can have significant impact on cultivation practices and crop production (Li et al.,

2005; Meilan et al., 2001; Robert et al., 1998). It can lead to the development of new crop varieties that are either earlier or later maturing, or do not even flower at all (Li et al., 2005; Silva et al., 2005).

In recent years, the developmental pathways that influence flower morphogenesis have become clearer with the help of genetic analysis, especially via the characterization of genes that are known to be directly involved in flower development in rapid-cycling accessions of *Arabidopsis* (Michaels et al., 2005). Extensive reviews of the mechanisms by which environmental and endogenous factors that regulate flowering time in *Arabidopsis*, dubbed as the 'Rosetta Stone' of flowering time, have been published (Putterill et al., 2004; Simpson and Dean, 2002; Simpson et al., 1999; Yanovsky and Kay, 2003).

A simplified model showing pathways that control flowering in *Arabidopsis* is illustrated in Figure 2. In these pathways, networks of regulatory genes are organized in a hierarchical fashion. Near the top of the hierarchy are regulatory genes conferring 'early' or 'late' flowering. Typically, these genes respond to environmental signals such as light and temperature.

It is assumed that the underlying variation in flowering mainly reflects adaptation to different environments to ensure reproductive success (El-Assal et al., 2001). Recent research that examines the regulatory genes and control of flowering of *Arabidopsis*, homologous genes in *Brassica* and other crops, and the evolution of the MADS-box gene family classification, which includes most of these genes, is reviewed below.

MADS-box genes

Most "molecular architects" of flower development belong to the family of MADS-box regulatory genes. Members of this family encode sequence-specific DNA-binding transcriptional activators that carry out a variety of developmental functions. In recent years, MADS-box genes that are expressed in vegetative tissues of agronomic crops and those that control reproductive

development have been explored with the aim of developing better cultivars (Aswath and Kim, 2005). The name MADS was coined based on the first four members of the gene family: *MCM1* (from *Sacchromyces cerevisiae*), *AGAMOUS* (from *Arabidopsis*), *DEFICIENS* (from *Anthirrhinum*) and *SRF* (a human protein). There are over 100 MADS-box sequences identified across eukaryotic species (Ng and Yanofsky, 2001). In the *Arabidopsis* genome, at least 47 types of MADS-box genes have been confirmed (Lawton-Rauh et al., 2000; Ratcliffe et al., 2001; Reichmann and Ratcliffe, 2000). MADS-box genes are similar to *HOX* genes in animals but are scattered in the plant genome, not organized in clusters. Proteins encoded by MADS-box genes are usually 240-260 amino acids long, 57 of which are in the conserved MADS domain (Purugganan, 1997).

The evolution of the MADS-box gene family has been postulated to have begun before the divergence of plants and animals (Alvarez-Buylla et al., 2000). The last common ancestor of extant plants, animals and fungi, which existed about one billion years ago, is hypothesized to have had two classes of MADS-box genes. From these two types, two different subfamilies of MADS-box genes, *ARG80-* and *MEF2*-like genes, evolved within the lineages leading to animals and fungi. A special class of MADS-box gene was gained somewhere in the lineage leading to green plants (Ng and Yanofsky, 2001). These produce proteins that follow an MIKC pattern – with conserved DNA-binding MADS domain (M) at the amino terminus, a weakly conserved intervening region (I), a moderately conserved keratin region (K), and a poorly conserved C-terminal region (C) (Theissen et al., 2000).

The MADS-box genes in green plants that are involved in floral organ development have been further classified into four floral homeotic gene groups, based on phylogenetic analyses – AP1/AGL9, PI, AP3, and AG. Alterations in these homeotic genes cause changes in patterns of floral development. The classification of the aforementioned homeotic genes into these four families was based on duplication events thought to predate the divergence of flowering plants and gymnosperms (Purugganan, 1997). However, research on ferns involving a higher number of MADS-box genes,

indicates that this divergence might have occurred more recently (Lawton-Rauh et al., 2000). There are more than 15 different members of the MADS-box family in the fern *Ceratopteris*. The last common ancestor of all higher angiosperms already had at least 10 different MIKC-type MADS-box genes, among them representatives of all the gene clades from which the floral homeotic genes have been recruited.

Research on the average DNA substitution rate of the plant MADS-box gene families reveals significant differences in the rates of non-synonymous nucleotide substitutions in different regions of the gene. The MADS domain, which is the DNA binding domain, was observed to have the lowest substitution rate, followed by the *K* region. The *C* region was determined to have the highest substitution rate, suggesting that this region is evolving at a near neutral rate (Purugganan et al., 1995). Overall, representative MADS-box regulatory genes were found to be evolving faster than structural genes based on relative non-synonymous to synonymous substitution rates (Barrier et al., 2001). In the Hawaiian silversword alliance (Asteraceae: Heliantheae-Madiinae), such accelerated rates of gene evolution likely resulted in rapid morphological diversifications (Barrier et al., 2001).

How MADS-box genes directly affected the actual evolution of certain plant parts is still not fully understood. At present, evidence of duplications followed by diversification has been found, as in the homeotic groups, and changes in expression patterns of regulatory genes were observed (Axelsson et al., 2001; Ng and Yanofsky, 2001). In *B. oleracea*, mutations in the MADS-box floral meristem identity genes, *CAULIFLOWER* and *APETALA1*, were associated with evolution of the domesticated cauliflower (Lowman and Purugganan, 1999; Purugganan et al., 2000). However, it is generally believed that the ability of MIKC proteins, coded for by MADS genes, to form heteromultimers (because of their K region) and generate complex protein networks is a factor influencing the rapid developmental diversification in plants (Kaufmann et al., 2005).

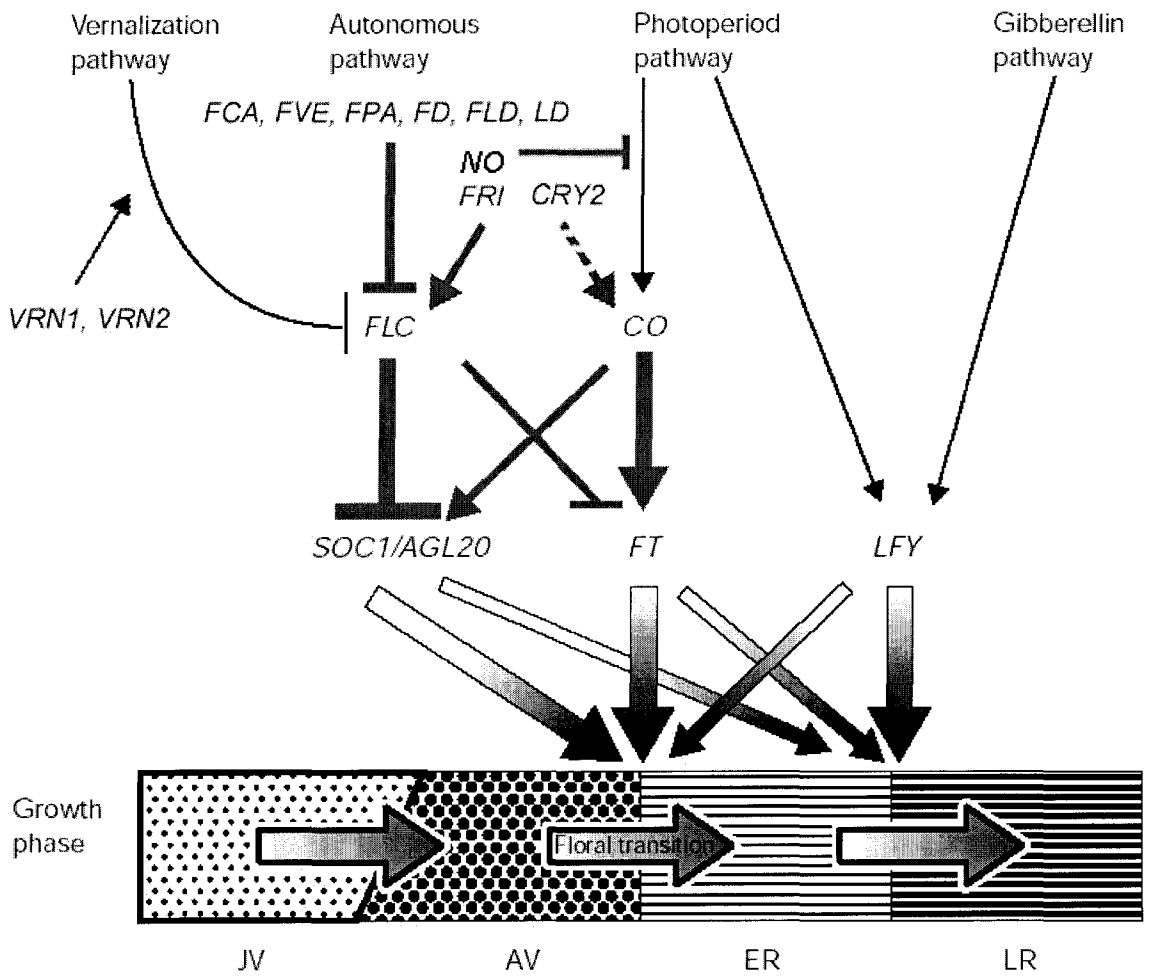


Figure 2. Model of floral transition in *Arabidopsis*. Modified figure from Araki (2001), with additional information from Mouradov et al. (2002), Quesada et al. (2005), Sheldon et al. (2000) and Simpson (2005). Abbreviations for growth phases are JV=juvenile vegetative, AV=adult vegetative, ER=early reproductive, LR=late reproductive.

Flowering-time genes

The identification of flowering-time genes in major crops was initiated by using QTL mapping and available information on analogous genes in *Arabidopsis* (Holland et al., 2002; Osborn et al., 1997). The following section reviews genes whose functions were determined to be directly related to flowering time. Most of the genes discussed are believed to be located high in the hierarchy of the various flowering pathways (Figure 2). Studies on target genes downstream of the pathway, such as *SOC1/AGL20* and *FT*, are also reviewed. These downstream genes, called 'integrators', were found to have both overlapping and independent functions in the determination of flowering time (Moon et al., 2005).

FLOWERING LOCUS C (FLC) and FLC-like genes. *FLC* is a known MADS-box gene that is closely related to the meristem and organ identity genes determining floral structure (Sheldon et al., 2000). *FLC* and *FLC*-like genes have been shown to be central to the control of the transition to flowering and the response to vernalization in *Arabidopsis* (Sheldon et al., 2000; Michaels et al., 2005). *FLC* represses the activity of the downstream genes *SOC1/AGL20* and *FT* (Figure 2), delaying flowering. A direct relationship exists between the activity of *FLC*-related genes and flowering time, with null and weak alleles of *FLC* effecting early flowering in *Arabidopsis* accessions (Werner et al., 2005). *FLC* encodes a MADS domain containing a transcription factor whose synthesis is promoted by the dominant *FRI* allele and inhibited by vernalization genes (*VRN1* and *VRN2*) through gene silencing by histone methylation (Amasino, 2005; Bastow et al., 2004; He and Amasino, 2004).

Approximately six members of the *FLC*-like gene family have been identified in *Arabidopsis* based on cDNA comparisons. The related genes, *F15O5.4, F15O5.3, MAF1, MXK3.30,* and *F15O5.2*, have been observed to have high sequence similarity with *FLC* and probably have functional

redundancy (Ratcliffe et al., 2001). *MAF1* (*MADS AFFECTING FLOWERING 1*) in particular, was also demonstrated to repress flowering, acting somewhere downstream of *FLC*. It has been subsequently named *FLOWERING LOCUS M* (*FLM*) (Scortecci et al., 2001). Analyses of nucleotide sequences of *MAF1* indicate an alternative splicing site resulting in a protein with a modified *I*-region and another which is 30 amino acids shorter (Ratcliffe et al., 2001). In addition to *MAF1*, other *FLC* paralogs were subsequently identified, including *MAF2, MAF3, MAF4 and MAF5.* These were also found to have varying effects on vernalization response in *Arabidopsis* (Ratcliffe et al., 2003)

Though *FLC* has been the focus of many studies concerned with regulatory control of flowering, naturally occurring variation in this gene in *Arabidopsis* was surveyed only recently and associated with clinal variation in flowering time (Stinchcombe et al., 2005). A certain allele class of *FLC* was found to have broader latitudinal distribution; the accessions in which it was observed also exhibited variation in vernalization response (Stinchcombe et al., 2005). Originally identified in *Arabidopsis, FLC*-related sequences were also found in the crucifers *B. oleracea*, *B. juncea, B. napus,* and *B. rapa* (Kole et al., 2001; Lin et al., 2005; Martynov and Khavkin, 2004; Tadege et al., 2001; Schrantz et al., 2002), and in other species, such as wheat (*Triticum aestivum* L.) and horseweed (*Conyza canadensis* L. (Cronq.)) (Rudnóy et al., 2002). In contrast, no *FLC* or *FLC*-like sequences have been identified in either expressed sequence tags or genome databases of pea (*Pisum*) and other legumes (Hecht et al., 2005) or in a species of wheat (*T. monococcum* L.) (Yan et al., 2004). Another gene similar to *FLC* might exist in legumes, as was reported in wheat (*T. monococcum)*, where the gene Zinc finger–CCT domain transcription factor (*ZCCT1*) *VRN2* functions similarly to *FLC* (Dubcovsky et al., 2006). It is hypothesized that the final regulatory steps of the flowering control pathway in *Arabidopsis* and the cereals evolved independently (Yan et al., 2004).

In *Brassica*, *FLC* is expected to have three copies in the diploid species and up to eight in the amphidiploids (Schrantz et al., 2002). These copy numbers were based on results of studies that suggest that the diploid *Brassica* species contains triplicated regions that correspond to the top of

Arabidopsis chromosome 5 (*At5*) (Osborn and Lukens, 2003). A fourth copy of *FLC* was later identified in *B. rapa* (Schrantz et al., 2002) and in *B. oleracea* (Udall, 2003). Other studies confirmed these copy numbers: one *FLC* homologous gene has been reported in *B. rapa* ssp. *pekinensis* (Li et al., 2005), two homologs have been identified in *B. oleracea* (Lin et al., 2005) and in *B. juncea* (Martynov and Khavkin, 2004), and eight copies in a resynthesized *B. napus* (Udall, 2003). The many copies of *FLC* in *B. napus* were found to have an effect by increasing the variation in flowering time in a dosage-dependent manner (Osborn, 2004). The sizes of *Brassica FLC,* from exon 2 to 7, range between 1000 to 2000 bp. Only exons 4 and 7 were reportedly variable in nucleotide length among the seven exons of the gene (Schrantz et al., 2002; Martynov and Khavkin, 2004). A survey of RFLP polymorphism in *FLC* in the cultivated *Brassica* indicated that variation within exons 4 and 7 can successfully distinguish among the species (Martynov and Khavkin, 2005b).

FRIGIDA (FRI) and FRI-like genes. *FRI* upregulates *FLC* activity; dominant *FRI* alleles are responsible for late flowering and vernalization requirements in natural accessions of *Arabidopsis* (Johanson et al., 2000; Michaels and Amasino, 2000). No study of *FRI* in *Brassica* was found in the literature.

FRI is a non-MADS-box gene; molecular analysis showed that early-flowering *Arabidopsis* accessions have alleles with one of two different deletions that disrupt its open-reading frame (Johanson et al., 2000). Rapid-cycling *Arabidopsis* plants - those that flower in the same season - may have evolved at least twice from late-flowering progenitors through loss-of-function *FRI* mutations. This loss-of-function may have resulted from multiple replacements and indels that rendered either of the two-coiled domains of the translated *FRI* protein inactive (Le Corre et al. 2002). In winter-annual *Arabidopsis*, several *FRI*-like sequences (*FRL1* and *FRL2*) were identified. These *FRI*-like genes function in the same way as *FRI* in upregulating *FLC*, causing a delay in flowering (Michaels et al., 2004).

A study of the sequence variation and structure of haplotypes surrounding the *FRI* gene region in *Arabidopsis* revealed that extensive and highly diverged haplotypes exist (Hagenblad and Nordborg, 2002). Variation in non-functional sequences was found to be higher than in functional sequences of *FRI*, indicative of a recent relaxation of selection pressure on the non-functional variants of the gene (Le Corre et al., 2002). It was demonstrated that about 70% of the variation in *Arabidopsis* flowering time could be attributed to *FRI* allelic variation (Shindo et al., 2005). Recent evidence indicates that a latitudinal cline in flowering time is caused by functional *FRI* alleles together with a certain allele class of *FLC* (Stinchcombe et al., 2004; Stinchcombe et al., 2005). Considering the entire gene, the extent of polymorphism in coding and non-coding regions has been observed to be nearly equivalent. However, there was also a departure from the neutral expectation for the whole gene and its coding regions. It is hypothesized that a purifying selection acted to maintain the functionality of the *FRI* protein, and maintain late flowering during the last glacial period when few habitats were accessible. When new environments or potential habitat became available after glaciation, selection for early flowering occurred, especially as suggested by non-synonymous polymorphisms in the first exon of *FRI* (Le Corre et al., 2002).

FLOWERING LOCUS CA (FCA). This gene evidently functions by repressing *FLC* activity, thus promoting flowering (Eckhardt, 2002). Mutants of this gene in *Arabidopsis* result in accumulation of *FLC* transcripts and, thus, late flowering. Plants that are *fca* mutants as well as those with dominant *FRI* behave as winter annuals. *FCA* is believed to interact with *FY*, a polyadenylation factor; together, they function to control polyadenylation site choice during processing of the *FCA* transcript (Henderson et al., 2005; Quesada et al., 2005).

The *Arabidopsis FCA* contains 20 introns. Alternative splicing between introns 3 and 13 produces four types of transcripts that are believed to play a role in regulation of meristem function in addition to influencing control of *FLC* and floral transition (Eckhardt, 2002; Macknight et al., 2002).

FCA genes have been identified in monocots, such as rice (*Oryza*) and ryegrass (*Lolium*), and were found to function similarly to the dicot flowering-time genes (Winichayakul et al., 2005). No *FCA* or *FCA*-like sequences in *Brassica* have been reported in the literature.

CONSTANS (CO) and CONSTANS-LIKE1 (COL1). *CO* and *COL1* genes are known to provide links between the control of the plant's circadian clock and flowering time (Yanovsky and Kay, 2002 and 2003). In *Arabidopsis,* transcript levels of *CO* were found to be regulated by the gene *EARLY FLOWERING 4* (*ELF4*), which is primarily responsible for photoperiod detection and circadian regulation. Loss-of-function mutations in *ELF4* result in early flowering under non-inductive photoperiods (Doyle et al., 2002). There were four copies of genes homologous to *Arabidopsis CO* isolated from *B. napus* (Robert et al., 1998). *CO* is directly involved in activating transcription of *LEAFY* (*LFY*), leading to floral transition (Figure 2). It encodes a zinc-finger protein, a transcriptional activator, and promotes flowering under long photoperiods. *CO* was proposed as the gene that should account for variation in flowering time in *B. nigra,* but no sequence variation was observed between early and late flowering plants (Österberg et al., 2002). In *B. napus,* one *CO* allele (*BnCOa1*) was identified that effects early flowering under any photoperiod (Robert et al., 1998).

The *Arabidopsis* gene, *CONSTANS-LIKE1* (*COL1*), shares sequence similarity with *CO* and is located 3.5 kb upstream of *CO*. *COL1* sequences have been identified in *B. nigra* and *B. rapa*; indel variation in the former species was found to be highly associated with flowering time (Martynov and Khavkin, 2005a; Österberg et al., 2002). Plants homozygous for the short-photoperiod allele flower earlier than those with a longer-photoperiod allele. Indel variation was also identified in *Arabidopsis* but may be absent in other closely related *Brassica* species, such as *B. oleracea* and *B. rapa.* This observation indicates that the indel was present in *B. nigra* before the split of the *Brassica* lineage (one to *B. nigra* and the other to *B. rapa* and *B. oleracea*) (Österberg et al., 2002). Neutrality tests however, failed to detect any evidence of selection though haplotype structures. This suggests that

either selection or recombination may have acted on the *Brassica* sample sequences (Lagercrantz et al., 2002). RFLP analysis also suggests a relationship between polymorphism in *CO*-like genes in *B. rapa* and geographic origin (Martynov and Khavkin, 2005a).

CRYPTOCHROME 2 (CRY2). *CRY2* is a member of a group of photoreceptors that are key regulators controlling various aspects of plant growth and development in response to the environment. *CRY2* might be involved in post-transcriptional activation of *CO* (Mouradov et al., 2002). El-Assal et al. (2001) described a natural phenotypic variant of *Arabidopsis* arising from a single amino-acid substitution in *CRY2* altering its normal protein function. The identified *CRY2* allele has single-nucleotide mutations in the structural gene sequences, and plants grown under short photoperiods were observed to flower early. No study of this gene in *Brassica* was found in the literature.

In general, the following observations from studies on flowering time genes indicate that:

1) Floral genes predate the origin of flowers, because of the presence of MADS-box genes in non-flowering plants and even animals (Alvarez-Buylla et al., 2000; Lawton-Rauh et al., 2000). From phylogenetic analysis, these genes were classified as members of previously identified floral homeotic groups orthologous to known angiosperm floral homeotic loci.

2) There is evidence that diversification in patterns of floral development and structures originate from molecular variation within species. Copy number and nucleotide variation in the flowering genes in *Brassica* may have contributed to wide variations in flowering time (Axelsson et al., 2001). In maize, molecular variation in the *teosinte-branched1* (*tb1*) gene was hypothesized to have resulted in differences in inflorescence architecture between maize and teosinte (Wang et al., 1999).

3) Changes in expression patterns and regulatory interactions between floral homeotic genes or changes in the targets of these genes determine the onset of flowering. The MADS-type floral homeotic genes have been prerequisites for the "invention" of flowers because they have been determined to predate both gymnosperms and angiosperms. On the other hand, it seems that the radiation of angiosperms and the evolution of flowering were based mainly on more subtle changes in expression patterns and regulatory interactions between already existing floral homeotic genes (Theissen et al., 2000). Regarding flowering time in both *Arabidopsis* and *Brassica,* variation in regulatory genes that lead to loss of function explains observed variation in flowering time. However, in legumes, the absence of these regulatory genes in the autonomous pathway of flowering suggests a different mechanism of flowering control, or the possibility that different genes are involved upstream of the modeled pathway.

Many researchers are still trying to unravel the structure of flowering time genes, how the function of many of these genes evolved, their genetic control, and the details and exact mechanism of floral development (Olsen et al. 2002; Ungerer et al., 2002). Numerous genes controlling flowering time have been identified together with an interacting network leading to flowering. However, limited information on naturally occurring variation in these flowering genes and on signaling pathways has been reported (Mouradov et al., 2002). As information becomes available, and as new genes from *Arabidopsis* or other plant species are identified, the current model of flowering development will be further refined. Studies on a wide range of plant species might also reveal interesting information, perhaps confirming whether the known flowering repressors follow a common molecular mechanism across taxa (Amasino, 2005). As demonstrated in legumes and temperate cereals, the model from *Arabidopsis* may not universally apply. In addition, research on perennial plants and woody species might reveal entirely different molecular mechanisms affecting the onset of flowering. The current model of flowering is entirely based on mutants of an annual life form of *Arabidopsis* (Battey et al., 2000; Simpson, 2005).

CHAPTER 2. CHARACTERIZATION OF FLOWERING TIME AND SSR MARKER ANALYSIS OF SPRING AND WINTER TYPE *BRASSICA NAPUS* L. GERMPLASM

A paper accepted for publication in Euphytica

Von Mark V. Cruz[1] , Richard Luhman[1]*, Laura F. Marek[1], Charlie L. Rife[2]**,

Randy C. Shoemaker[1,3], E. Charles Brummer[1] & Candice A.C. Gardner***[1,4]

Abstract

Flowering dates and life forms of all available *Brassica napus* accessions conserved at the North Central Regional Plant Introduction Station (NCRPIS) were characterized, and a survey of molecular variation was conducted by using simple sequence repeats (SSR) in order to support better management of accessions with diverse life forms. To characterize flowering phenology, 598 *B. napus* accessions from the NCRPIS collection were planted in Iowa and Kansas field sites together

---

[1]Department of Agronomy, Iowa State University, Ames, IA 50011

[2]Department of Agronomy, Kansas State University, Manhattan, KS 66506

[3]USDA-ARS Corn Insect and Crop Genetics Research Unit, Ames, IA 50011

[4]USDA-ARS Plant Introduction Research Unit and North Central Regional Plant Introduction Station, Ames, IA 50011

* *present address: Monsanto Co., Saint Louis, MO 63167*

** *present address: Blue Sun Biodiesel, Westminster, CO 80234*

*** *author for correspondence: email: gardnerc@iastate.edu*

with a current commercial cultivar and observed for days to flowering (first, 50% and 100% flowering) in 2003. Days from planting to 50% flowering ranged from 34 to 83 in Iowa and from 53 to 89 in Kansas. The mean accumulated growing degree days (GDD) to 50% flowering were 1997 in Iowa, and 2106 in Kansas. Between locations, the correlation in flowering time (r=0.42) and the correlation in computed growing degree days (r=0.40) were both significant. Differences in flowering-time rank were observed for several accessions. Accessions that failed to flower in Iowa in a single growing season comprised 28.5% of the accessions; of the flowering accessions, 100% plant flowering was not always achieved. Accessions were grouped according to flowering time. A stratified sample of 50 accessions was selected from these groups, including 10 non-flowering and 40 flowering accessions of diverse geographic origins and phenological variation. The flowering time observed in the sampled accessions when grown in the greenhouse were found to be significantly correlated to the flowering time observed in the field locations in Iowa (r=0.79) and Kansas (r=0.49). Thirty SSR markers, selected across 18 *Brassica* linkage groups from BrassicaDB, and three derived from *Brassica* expressed sequence tags (ESTs) were scored in the stratified sample. An average of three bands per SSR primer pair was observed. Associations of SSR marker fragments with the life forms were determined. Analysis of molecular variation by using cluster analysis and ordination resulted in recognizable, distinct groups of annual and biennial life-form types, which may have direct applications for planning and management of future seed regenerations.

Keywords: *Brassica napus*, diversity, genebank, microsatellites, phenology, rapeseed

Introduction

Brassica napus is an amphidiploid (AACC genome, 2n=38) that is believed to have arisen from interspecific hybridization between the diploid species *B. rapa* L. (syn. *B. campestris*; AA genome, 2n=20) and *B. oleracea* L. (CC genome, 2n=19) (Sauer, 1993; Gómez-Campo, 1999). *Brassica napus* has a short evolutionary history since it is presumed to have arisen from cultivation, and no true wild forms have been found (Lackey, 1996; Gómez-Campo and Prakash, 1999). This species is believed to have originated in the Mediterranean region of southwestern Europe where native populations of *B. rapa* and *B. oleracea* overlap. *Brassica napus* populations can be classified into annual or biennial forms. The annual forms are called spring rapes and the biennial forms are called winter rapes (McNaughton, 1995). Spring rapes are sown and harvested in the same season, while winter rapes are sown in the fall and require vernalization before flowering in the next season (Butruille et al., 1999). It is believed that annual and winter type *B. napus* are distinct groups; intercrossing between these types is still not common even in contemporary breeding programs (Diers and Osborn, 1994).

Extensive genetic resource collections of *Brassica* exist to support both public and private breeding programs (Boukema and van Hintum, 1999). The USDA-ARS North Central Plant Introduction (NCRPIS) in Ames, Iowa has been mandated in the U.S. to curate germplasm of *B. napus* along with more than twenty other species of *Brassica* and their wild and weedy relatives. Both annual and biennial populations are represented among the accessions conserved by the NCRPIS, and possibly some of mixed life forms. However, not every accession is linked to passport or characterization data describing the type of life form. Information on an accession's life form prior to regeneration facilitates more efficient resource management by identifying those that require vernalization, a time and labor-intensive process that requires specialized facilities. If accessions consist of mixed life forms, regeneration methods can be modified to preserve their genetic profiles.

To date, all types have been subjected to vernalization treatment and then transplanted to the field as part of standard regeneration procedures. More accurate information on the flowering data of annual types can be collected without vernalization, since exposure to cold treatment may lead to precocious flowering (Friend, 1985; Sovero, 1993). At present, the available *B. napus* flowering data on the ARS-GRIN database (http://www.ars-grin.gov/npgs) indicate 'days to flower' which include the number of days from the time accessions are germinated in blotter boxes, vernalized, and transplanted to the field. To complement the existing data in the GRIN database, we conducted this study and gathered flowering data on annual accessions without vernalization. The specific objectives are to (i) determine life forms of all available *B. napus* accessions by using morphological and phenological characterization, (ii) survey the molecular genetic variation of representative accessions, and (iii) identify life-form specific markers to possibly help curators in screening new accessions quickly without waiting for actual flowering.

To address the last two of the aforementioned objectives, simple sequence repeats (SSR) or microsatellite markers were used. SSRs are fairly robust, have good reproducibility, and are more cost-effective compared to other marker systems (Farooq and Azam, 2002). It is anticipated that the GRIN database will include SSR profiles of germplasm accessions; some crops already have SSR data. In *Brassica*, SSRs have been valuable tools, and numerous SSR primer sequences are already publicly available (Snowdon and Friedt, 2004). SSRs have been used in *Brassica* research for such topics as seed-coat color mapping (Padmaja et al., 2005), varietal identification (Tonguç and Griffiths, 2004), and analysis of variation in plant populations and germplasm collections (Raybould et. al., 1999; Westman and Kresovich, 1999). It was also reported that a set of SSRs was able to identify groups of annual and biennial types of *B. napus* elite lines (Plieske and Struss, 2001; Tommasini et al., 2003). Here we test whether it is also possible to identify life-form types by using diverse, heterogeneous germplasm collections of *B. napus*.

Methodology

Planting and characterization of flowering time. In May 2003, 598 accessions of *B. napus* originating from 28 countries were planted at the NCRPIS farm in Ames, Iowa together with 'Hyola 401', a commercial spring cultivar. In Ames, the accessions were randomized and planted in three replicates by direct seeding in 7-meter, single-row plots in a 0.78 ha field located at 42°00'29.61"N, 93°39'48.82"W. In April 2003, these same 598 accessions were also planted in 7-meter, single-row plots without replication (due to resource limitations) at Kansas State University (KSU), Manhattan, Kansas. Data gathered in Iowa included (1) date of 50% germination, (2) date of first flowering (when the first plant in the plot flowered), (3) date of 50% flowering (when 50% of the plants in the plot flowered), and (4) date of 100% flowering. Data gathered in Kansas included date of 50% flowering. Following characterization during the 2003 field season, 50 accessions representing the range of flowering times observed as well as diverse geographic sources were selected (Table 1). These accessions were comprised of 10 accessions of non-flowering types (putative biennials) and 40 of flowering types (putative annuals). Selection was done by hierarchical clustering of the flowering time and subsequent selection from each of the resulting eight clusters based on geographic origin. The selected accessions included cultivars and breeding lines. All of the selected representative accessions from among the flowering types attained at least 50% flowering in the field in both locations.

In March 2004, seeds from the original seed lot of the selected accessions were sown in 10 cm square plastic pots in the greenhouse containing Sunshine Mix no.1 (Sun Gro Horticulture, Bellevue, WA). Twenty-four plants for each accession were grown under a 16-hour photoperiod to verify the life form, determine correlations between flowering time in the greenhouse and the field, and to obtain leaf tissues for molecular-marker analysis. Mean greenhouse temperature and illuminance were recorded at 26°C and 3.9 kilolux, respectively, by HOBO® H08-004-02 data loggers

(Onset Computer Corporation, Bourne, MA). The data loggers were positioned at the leaf canopy level.

Plant material and DNA extraction. Leaf tissues were harvested when the plants in the greenhouse were at the 3-4 leaf stage, when the leaves contain relatively low amounts of polysaccharides (Hyam, 1998). Twelve to fifteen plants were sampled per accession. The leaves were lyophilized by freeze-drying and then stored in a -80°C freezer until use. Genomic DNA extraction was performed by using 1 g bulked tissue obtained from equal weights of freeze-dried leaf samples from individual plants within an accession. The bulked tissues were placed in 50 ml screw-cap polypropylene tubes, each containing approximately 1.5 ml glass beads. The tissue samples were ground to a powder by agitating the tubes on a paint shaker for 90 seconds. DNA was isolated using the cetyltrimethyl ammonium bromide (CTAB) method. The extracted DNA was resuspended in 1X TE (1 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) with 3 μl of 10 mg/ml RNAseA then incubated at 30°C for 30 minutes and placed on a LabQuake® shaker (Barnstead Intl., Dubuque, IA) at 4°C overnight. The samples were stored at -20°C until use.

PCR amplification of microsatellites. Microsatellite primers were synthesized by Integrated DNA Technologies (IDT, Coralville, IA) based on published primer sequences from Uzunova and Ecke (1999), Westman and Kresovich (1999) and the BBSRC BrassicaDB (http://brassica.bbsrc.ac.uk). The 33 microsatellite loci are presented in Table 2. To ensure representation of several linkage groups, the microsatellites from BBSRC were selected based on indicated linkage map locations in BrassicaDB and from Lowe et al. (2004). These microsatellites were developed mostly from genomic libraries (Lowe et al., 2004). Three primer pairs derived from *Brassica* EST sequences were provided by Dr. Andrew Salywon (USDA-Agricultural Research Service, U.S. Water Conservation Laboratory, Phoenix, AZ) (see Salywon et al., 2004 for detailed descriptions).

PCR was performed in 96-well Microseal™ polypropylene microplates (Bio-Rad Lab. Inc., Hercules, CA) with each sample well containing 1 μl of genomic DNA (50 ng/μl), 8 μl of sterile ddH2O, 1 μl 10X PCR buffer (200 mM Tris-HCl, pH 8.4, 500 mM KCl), 0.5 μl dNTPs (2 mM), 0.3 μl MgCl (50 mM), 0.1 μl of primers (50 μM), and 0.05 μl *Taq* polymerase (5 U/μl) (Invitrogen Corp., Carslbad, CA). A negative control, a sample from a commercial hybrid (Hyola 401), and samples from three rapid-cycling *Brassica* accessions (*B. oleracea* TO 1000 DH3, *B. rapa* IMB 218 DH3, *B. napus* EL 6400 A) from Dr. J. Chris Pires (Univ. Missouri, Columbia, MO) were included in each microplate. Thermal cycling was done by using DNA Engine® (PTC-200™) thermal cyclers (Bio-Rad Lab. Inc.) under the following conditions: 94°C for 2 min, then followed by 35 cycles of amplification at 94°C for 30 sec, 55.5°C for 30 sec, 72°C for 30 sec, followed by a final extension at 72°C for 4 min. Different annealing temperatures were used on two of the EST-derived primers, 53.4°C for EST 1 and 59.5°C for EST 2. All reactions with no initial amplifications were repeated to confirm the results. The PCR products were separated in a 4.0% agarose gel prepared in 1X TAE (40 mM Tris-acetate, 1 mM EDTA, pH 8.3) with incorporated ethidium bromide (0.46 μg/ml). Amplification products were visualized with a UV light box (254 nm wavelength) and photographed on a digital gel-documentation system. The bands were scored as present (1) or absent (0) and recorded with reference to the molecular weight markers. They were then treated for analysis as if they were dominant markers, since it was not possible to calculate gene frequencies as a result of bulking 10 to 15 plants per accession.

Data analysis. Five accessions were not included in subsequent analyses after determining that they were not likely to belong to the *B. napus* species; two accessions are probable members of *B. rapa* (PI 286418, Ames 21490), two of *B. juncea* (L.) Czern. (Ames 19197, Ames 24222) and one of *B. oleracea* (PI 357374). Likewise, changes in country of origin of two accessions (see note Table 1)

were used in the computation of geographic distances. Descriptive statistics and preliminary cluster analysis of flowering data were obtained from JMP® version 5.1.2 software (SAS Institute, Cary, NC). Growing Degree Days (GDD) were computed for all flowering accessions by using the formula GDD=(*MinT*+*MaxT*)/2 – *Base T*, where *MinT* is the lowest temperature of the day and *MaxT* is the highest temperature of the day; Temperature thresholds for *MinT* and *MaxT* were set to 0°C and 30°C, respectively. *Base T* is the temperature below which no development occurs and was set to 0°C, based on recent research which indicated that it was more accurate for *B. napus* plant development than is the typical base temperature of 5°C (Thomas, 2003).

Because of bulking, analysis of SSR bands was done following a shared-alleles method. Nei and Li (1979) distance was computed using the formula $d_D=1-(2v_{ij}/2v_{ij}+w_{ij}+x_{ij})$, where v_{ij} is the number of bands in common between both accessions; w_{ij} is the number of bands present in the *i*th accession and absent in the *j*th accession; x_{ij} is the number of bands absent in the *i*th accession and present in the *j*th accession (Reif et al., 2005). Cluster analysis was performed by using the neighbor-joining method algorithm in the NTSYS-pc version 2.20e software package (Rohlf, 2005). A Chi-square test for independence was done to test the hypothesis that observed band frequencies were independent of life form. Associations between band presence and flowering time were made using the non-parametric Mann-Whitney-U test (equivalent to Wilcoxon rank sum test) (Gebhart et al., 2004). Mantel tests were performed to determine the correlation between distance matrices (Koenig, 1999) and analysis of molecular variance (AMOVA) to test for genetic differentiation. The last two aforementioned tests, as well as calculations of genetic distances based on shared alleles (Maguire et al., 2002) and geographic distances (computed using latitude-longitude coordinates of the accessions' origin), and principal coordinate analysis (PCA) were done by using the GenAlEx v.6 software package (Peakall and Smouse, 2006).

Results

Characterization of flowering time

One hundred sixty nine accessions, which represent 28.5% of the *B. napus* collection, did not flower in Iowa. Four hundred twenty five accessions (71.5%) reached first flowering, 279 accessions (47.0%) continued to 50% flowering, and 198 accessions (33.3%) achieved 100% flowering. In the Ames location, the mean number of days after planting (DAP) to 50% germination was 13 days and the mean DAP for first flowering was 49. For the 279 accessions that continued to the 50% flowering stage, the mean number of DAP was 55. The remainder did not reach that stage by 100 DAP when scoring ended. For the 198 accessions that achieved 100% flowering, the mean number of DAP was also 55. This reflects the phenomenon that those accessions that most quickly reached the 50% flowering stage also generally achieved 100% flowering first. For those accessions that achieved 100% flowering, on average, it took 7 days from first to 50% flowering, and 3 days from 50% flowering to 100% flowering. PI 537302 was the earliest flowering accession, attaining 50% flowering at 34 DAP, while the very late flowering accessions PI 458940, PI 469770, and PI 470075 attained 50% flowering at 83 DAP. The plants of about 32% of the flowering accessions flowered synchronously (Figure 1A).

In Kansas, 50% flowering was observed in 223 accessions with a mean of 74 DAP. In general, more *B. napus* accessions took longer to flower when planted in Kansas (Figure 1). Correlation between the 50% flowering dates observed in Iowa and Kansas was low (r=0.42) but significant ($p<0.0001$). There was a mean difference of four days between 50% flowering in the two locations. The commercial cultivar Hyola 401 is the earliest to attain 50% flowering followed by Ames 26654 at 53 DAP and 54 DAP, respectively. Sixteen accessions reached 50% flowering at 89 DAP.

The flowering times observed in the greenhouse for 50 selected representative accessions were found to be correlated to flowering times observed in the field in both Iowa ($r=0.79$, $p<0.0001$) and Kansas ($r=0.43$, $p=0.008$). However, the mean DAP to flowering in the greenhouse were 62 to first flowering and 92 to 50% flowering, indicating that temperature and photoperiod differences may have delayed the development of flowering in a spring greenhouse relative to summer field conditions.

Accumulated GDDs to flower were computed for all flowering accessions. GDDs can better describe the conditions required to reach physiological developmental stages, such as flowering, than 'days to maturity' alone (Eckert, 2005). For data from Iowa, the accumulated GDDs were determined as follows: 207 ± 2 (mean, s.e.) for 50% germination, 1008 ± 9 for first flowering, 1997 ± 29 for 50% flowering, and 2021 ± 28 for 100% flowering. For data from Kansas, the accumulated GDD for 50% flowering was 2106 ± 22. Correlation between the accumulated GDDs to 50% flowering in the two locations is low but significant ($r=0.40$, $p<0.0001$). A scatter plot matrix of the GDDs in the two locations is shown in Figure 2. The estimates of GDDs in the two locations used in this study exceed the means and range of GDDs for each growth stage compiled from research data of Agriculture and Agri-Food Canada (AAFC) for *B. napus* (Thomas, 2003). The large mean values might have been influenced by the predominance of mid- and late-flowering accessions in the collection relative to those cultivated in Canada.

Microsatellite analysis

A total of 98 bands and an average of 3 bands per primer pair were observed based on all 33 SSRs and bulked samples from 50 selected accessions. The primer pairs all showed amplification, but band presence in EST1 was observed in only one accession. No polymorphism was observed in SSR 59A1. Single-band products were observed in four of the SSR primer pairs, while the remaining

primer pairs amplified two to eight products (Table 2). The number and size range of observed bands amplified from the bulked tissues were consistent with other studies (Allender, 2004; Tommasini et al., 2003). Across the 50 selected accessions, 71.9% of bands showed presence/absence polymorphism. The rate of polymorphism was higher among the flowering types, 82.7%, than for the non-flowering types, 61.2%. A chi-square test for independence indicated that there were significant differences in band frequencies between the non-flowering and flowering types ($p<0.0001$). Seventeen bands from 11 SSRs were observed in only one type of life form; twelve bands were observed solely in flowering types and five solely in the non-flowering types. These bands were from SSRs 25A, EST3, MR176, Na10-B08, Na12-A08, Na12-C07, Na14-C12, Ol10-F11, Ol12-G04, Ra2-E07, and Ra3-H10.

The observed frequencies of the unique bands in the non-flowering types were very low. However, results of Mann-Whitney-U tests indicated that band presence in EST2, Na10-B08, Ol11-C02, Ra2-E07, and Ra2-F11 exhibited significant associations with flowering time (Table 3).

A plot of the first two principal coordinates (PCO) derived from computed genetic distances from SSR-marker profiles allowed the visualization of relationships among accessions (Figure 3). Axis 1 explains 25.81% of the observed variation and axis 2, 18.35%. The non-flowering types grouped in quadrant IV, as did known early-flowering cultivars in quadrant I. The pattern observed in the PCO plot closely resembles relationships among accessions when cluster analysis was performed. The dendrogram separated flowering and non-flowering types, and early-flowering accessions mostly grouped together (Figure 4). Among the non-flowering types, only 539N (PI 535866, Silesia) was not within the same cluster. No distinct clustering of mid- or late-flowering accessions was observed. The computed, mean distance among the non-flowering types was 0.23, and among the flowering types it was 0.28. The smallest distance observed was 0.06 between 31 (Ames 19204) and 32 (Ames 19205); both originated from the former Soviet Union. Other accessions with the same country of origin also grouped together: 113 (PI 458940) and 114 (PI 458941) from Japan, 119 (PI 458948) and 126 (PI

458955) from Germany, and 62 (PI 311727) and 574 (Ames 22548) from Poland. Accessions with the same variety name also grouped together: 62 (PI 311727) and 574 (Ames 22548) - 'Bronowski', and 327 (PI 469911), 340 (PI 469924), and 349 (PI 469933) - 'Mokpo'.

The dendrogram indicates greater similarity between the rapid cycling *B. oleracea* accession and representatives of *B. napus* as compared to the *B. rapa* control. This observation supports a phylogeny proposed by Pradhan et al. (1992) among the *Brassica* species that was derived from analyses of cpDNA and mtDNA polymorphisms. 'Hyola 401' is in a cluster with the early-flowering accessions and known commercial canola cultivars, including 36 (Ames 2665, 'Westar'), 21 (Ames 15939, 'Comet'), and 108 (PI 458930, 'Oro'). Pedigree information was not available in the passport data of the canola cultivars, but their clustering may be due to highly similar genetic background. It has been suggested that there is a single common origin for most of the oilseed *B. napus* cultivars from results of RFLP analysis (Song and Osborn, 1992).

Results of Mantel tests indicated that the computed genetic distances are weakly correlated with geographic distances between the countries of origin ($r=0.35$, $p<0.001$). The significance of these correlations indicates that accessions with proximate geographic origin are somewhat more likely to have similar genetic profiles. Results of analysis of molecular variance (AMOVA) among different levels are presented in Table 4. The population genetic differentiation based on the binary data was given as Φpt, which is analogous to *Fst* (Peakall and Smouse, 2005). *Fst* values range from 0, when the subpopulations are identical in allele frequencies, to 0.5, if they are fixed for different alleles. The computed values considering life forms ($\Phi pt = 0.11$) and geographical origin ($\Phi pt = 0.02$ to 0.12) indicate that there is low to moderate genetic differentiation between the specified groupings. Grouping by life forms and by geographical regions explain 11% and 6% of the molecular variation, respectively. Much of the differentiation remained within the groupings in all comparisons with explained variation ranging from 88% to 98%. The results of the AMOVA analysis imply that there is some gene flow between life forms and more extensive gene exchange among geographic regions.

Diers and Osborn (1994) indicated that intercrossing between annual and winter type *B. napus* was not common even in breeding programs at that time. The data obtained from clustering and principal coordinate analyses as discussed above support Diers and Osborn's statements that these *B. napus* life forms are distinct groups.

Discussion

In the conservation of plant genetic resources, the availability of characterization data help germplasm users identify the accessions of interest and also provide plant breeders initial data for use in crop improvement programs. In this study, a survey of molecular variation was conducted and additional information was obtained to determine the life form of the *B. napus* accessions in the NCRPIS collection. The observed differences in the flowering time of annual accessions grown in two locations gave some information on the effect of environment on the flowering of *B. napus*. The flowering-time differences might have been caused by the varying responses of the genotypes to temperature and photoperiod (Friend, 1985). Such differences suggest that it is desirable to conduct characterization and evaluation trials for *Brassica* germplasm in more than one location, with distinct temperature and photoperiod regimens if resources permit. In the model species *Arabidopsis*, photoperiod and temperature affect the variation in flowering time by targeting expression of the *CONSTANS* and *FLOWERING LOCUS C* (*FLC*) genes located in different pathways. The major gene in the *Arabidopsis* photoperiod pathway is the *CONSTANS* gene, while in the autonomous and vernalization pathways the major gene is *FLC* (Simpson et al., 1999). Research progress in *Arabidopsis* has enabled the identification and cloning of related genes in *Brassica* that are known to directly influence the onset of flowering (Lagercrantz et al., 2002; Martynov and Khavkin, 2004, Osborn and Lukens, 2003). In *Brassica*, *CONSTANS* has been proposed to have the greatest influence in the flowering time variation in *B. nigra* and *B. oleracea* (Bouhon et al., 1998; Lagercrantz et al.,

1996). In *B. napus,* studies suggest that it is the *FLC* gene that influences flowering time in this species (Osborn and Lukens, 2003). Exposure to vernalization treatment diminishes the effect of late-flowering alleles of *FLC* (Osborn and Lukens, 2003). We have also examined and compared the *FLC* sequence variation between spring and winter types for the same set of accessions used in this study (Cruz, 2006),

In several accessions of *B. napus* germplasm regenerated in this study, we observed differences in flowering time within accessions that could be a concern for germplasm managers. The heterogeneity of flowering within accessions during regeneration has the potential to influence changes in allelic frequency and genetic profile over time, through assortative mating and possible selection against infrequent phenological phenotypes or those that do not flower under local conditions. Significant shifts in phenotype frequencies have been documented to occur in *B. napus* with just one cycle of germplasm regeneration (Diaz et al., 1997). In the present study, more than half of all flowering accessions did not progress to complete flowering. Only plants that flowered will be represented in the regenerated sample, resulting in possible genetic shifts. Non-flowering of spring types and of vernalized winter *Brassica* has also been documented and attributed to the effects of high temperatures, which often cause devernalization (Dahanayake and Galwey, 1998). If needed, the floral induction can be accomplished through the application of gibberellins at the *B. napus* rosette stage (Dahanayake and Galwey, 1999).

We have determined that it is possible to classify diverse accessions of *B. napus* into life-form types by using SSRs. This finding corroborates past results of studies in breeding lines and elite cultivars of winter and spring types (Charters et al., 1996; Lombard et al., 1999; Plieske and Struss, 2001; Tommasini et al., 2003). Our study further demonstrates that a limited set of SSRs can generate sufficient variability to distinguish among the life-form types in our set of germplasm. The data generated using SSRs may also prove useful in breeding programs. The results of the SSR analysis may have applications in screening new sets of *B. napus* accessions without life-form data or possibly

in identifying early or late-flowering accessions. The usefulness of marker-trait associations in assessing the genetic potential in germplasm collections has been previously studied in crops such as alfalfa (Skinner et al., 2000), cotton (Abdurakmonov et al., 2005), potato (Gebhart et al. 2004), and rice (Kadirvel and Gunathilagaraj, 2003). In *B. napus*, it has been demonstrated that molecular markers can be used in identifying genotypes with certain linolenic and erucic acid levels (Rajcan et al., 1999) and in predicting hybrid performance, as suggested by significant correlations between agronomic traits, such as plant height and seed yield, and genetic distances from markers (Riaz et al., 2001; Yu et al., 2005).

Summary and Conclusions

Characterization of flowering time in the field enabled the identification of life forms of genebank accessions of *B. napus* that had not been previously regenerated. The flowering dates obtained will be made available to complement existing data on the ARS-GRIN database. Our flowering dates should better represent a true phenology for the annual types, since the seedlings were not vernalized. An observed shift in flowering time in several accessions when planted in different locations was observed; this indicates that future flowering-data characterizations should be conducted by using multiple sites and interpreted accordingly.

Clustering based on marker profiles derived from microsatellites segregated groups of life forms and also identified probable duplicates (accessions with the same name). Results of the ordination analysis and AMOVA revealed that diversity 'within' life forms is greater than 'between' types of life forms in our selected set of representative accessions, which is in agreement with results from the clustering analysis. Significant, but relatively weak associations between derived genetic distances and geographic origins and between derived genetic distances and life forms were found. Seventeen bands from twelve SSR loci were observed to be uniquely present within accessions

exhibiting a particular life form, but often these occurred with low frequencies. Additional evaluation of 11 SSR loci identified significant associations of alleles with flowering time. Additional investigation is needed to determine if these are in non-coding genomic regions, proximal to or within genomic regions that control vernalization and flowering response, or in linkage disequilibrium with such regions.

Acknowledgements

This is a journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 1018, and was supported by Hatch Act and the State of Iowa. We would like to thank T.C. Osborn (Seminis) for advice during the initial phases of the study, J.C. Pires (Univ. Missouri) for the DNA from rapid cycling *Brassica* lines, A. Salywon (USDA-ARS USWCL) for the EST primers, and the dedicated staff of USDA-ARS NCRPIS and the USDA-ARS Corn Insect and Crop Genetics Research Unit for help in seed preparation, planting, field maintenance and technical assistance in the laboratory. We are also grateful for the Fulbright Program of the U.S. Department of State's support of VMVC during this study. Mention of commercial brand names in this paper does not constitute an endorsement of any product by the U.S. Department of Agriculture or cooperating agencies.

References

Abdurakhmonov, I.Y., A.A. Abdullaev, S. Saha, Z.T. Buriev, D. Arslanov, Z. Kuryazov, G.T. Mavlonov, S.M. Rizaeva, U.K. Reddy, J.N. Jenkins, A. Abdullaev & A. Abdukarimov. 2005. Simple sequence repeat marker associated with a natural leaf defoliation trait in tetraploid cotton. J. Heredity. 96:644-653.

Allender C.J. 2004. Evaluation of 48 public domain *Brassica* microsatellites. http://www.brassica.info/ssr/SSRinfosheet.xls (cited 11-4-05).

Boukema, I.W. & T.J.L. van Hintum. 1999. Genetic resources. In: Biology of *Brassica* Coenospecies. (Developments in Plant Genetics and Breeding, vol. 4). C. Gómez-Campo (ed.). Elsevier Science, Amsterdam, The Netherlands. pp. 461-479.

Bouhon.E.J.R., L.D. Ramsay, J.A. Craft, A.E. Arthur, D.F. Marshall, D.J. Lydiate & M.J. Kearsey. 1998. The association of flowering time quantitative loci with duplicated regions and candidate loci in *Brassica oleracea*. Genetics 150:393-401.

Butruille, D.V., R.P. Guries & T.C. Osborn. 1999. Linkage analysis of molecular markers and quantitative trait loci in populations of inbred backcross lines of *Brassicas napus* L. Genetics 153:949-964.

Charters, Y.M., A. Robertson, M.J. Wilkinson & G. Ramsay. 1996. PCR analysis of oilseed rape cultivars (*Brassica napus* L. ssp. *oleifera*) using 5'-anchored simple sequence repeat (SSR) primers. Theor. Appl. Genet. 92:442-447.

Cruz, V.M.V. 2006. Molecular genetic analysis of oilseed *Brassica*: Determination of life forms and germplasm management strategies by using microsatellite markers and *FLOWERING LOCUS C* gene sequences. Ph.D. dissertation. Iowa State University, Ames, Iowa.

Dahanayake, S.R. & N.W. Galwey. 1998. Effects of interactions between low and high temperature treatments on flowering of spring rape. Ann. Bot. 81:609-617.

Dahanayake, S.R. & N.W. Galwey. 1999. Effects of interactions between low-temperature treatments, gibberellins (GA_3) and photoperiod on flowering and stem height of spring rape. Ann. Bot 84:321-327.

Diaz, O., M. Gustafsson, & D. Astley. 1997. Effect of regeneration procedures on genetic diversity in *Brassica napus* and *B. rapa* as estimated by isozyme analysis. Genet. Res. Crop. Evol. 44: 523–532.

Diers, B.W. & T.C Osborn. 1994. Genetic diversity of oilseed *Brassica napus* germplasm based on restriction fragment length polymorphism. Theor. Appl. Genet. 88:662-668.

Eckert, D.J. 2005. Growing Degree Days as a Method of Rating the Maturity of Corn Hybrids. Ohio State University Extension. Agronomy Fact Sheet no. AGF-101-90. http://ohioline.osu.edu/agf-fact/0101.html (cited 11-4-05).

Farooq S. & F. Azam. 2002. Molecular markers in plant breeding-II. Some pre-requisites for use. Pakistan J. Biol. Sci. 5:1141-1147.

Friend, D.J.C.1985. *Brassica.* In: CRC Handbook of Flowering (Vol. II). A.H. Halevy (ed.) CRC Press Inc., Boca Raton, FL. pp. 48-77.

Gebhart, C., A. Ballvora, B. Walkemeier, P. Oberhagemann & K. Schuler. 2004. Assessing genetic potential in germplasm collections of crop plants by marker-trait association: A case study for potatoes with qualitative variation of resistance to late blight and maturity type. Molecular Breeding 13:93-102.

Gómez-Campo, C. 1999. Taxonomy. In: Biology of *Brassica* Coenospecies. (Developments in Plant Genetics and Breeding, vol. 4). C. Gómez-Campo (ed.). Elsevier Science, Amsterdam, The Netherlands. pp. 3-32.

Gómez-Campo, C. & S. Prakash. 1999. Origin and domestication. In: Biology of *Brassica* Coenospecies. (Developments in Plant Genetics and Breeding, vol. 4). C. Gómez-Campo (ed.). Elsevier Science, Amsterdam, The Netherlands. pp. 33-58.

Hyam, R. 1998. Field collection: Plants. In: Molecular Tools for Screening Biodiversity. A. Karp, P.G. Isaac, and D.S. Ingram (eds.). Chapman & Hall, London. pp. 49-50.

Kadirvel, P. & K. Gunathilagaraj. 2003. Detection of simple sequence repeat markers associated with resistance to whitebacked planthopper, *Sogatella furcifera* (Horvath), in rice. International Rice Research Notes 28(2):22-23.

Koenig, W.D. 1999. Spatial autocorrelation of ecological phenomena. TREE 14:22-26.

Lackey, J. 1996. Biology of rapeseed. USDA - Animal and Plant Health Inspection Service (APHIS). http://www.aphis.usda.gov/biotech/rapeseed.html (cited 11-4-05).

Lagercrantz, U., M.K. Österberg, & M. Lascoux. 2002. Sequence variation and haplotype structure at the putative flowering-time locus *COL1* of *Brassica nigra.* Mol. Biol. Evol. 19:1474-1482

Lombard, V., C.P. Baril, P. Dubreuil, F. Blouet, & D. Zhang. 1999. Potential use of AFLP markers for the distinction of rapeseed cultivars. In: New Horizons for an Old Crop. Proceedings of the 10th International Rapeseed Congress. Canberra, Australia. http://www.regional.org.au/au/gcirc/4/587.htm (cited 11-4-05).

Lowe, A.J., C. Moule, M. Trick & K.J. Edwards. 2004. Efficient large-scale development of microsatellites for marker and mapping applications in *Brassica* crop species. Theor. Appl. Genet. 108:1103-1112.

Maguire T.L., R. Peakall & P. Saenger. 2002. Comparative analysis of genetic diversity in the mangrove species *Avicennia marina* (Forsk.) Vierh. (Avicenniaceae) detected by AFLPs and SSRs. Theor. Appl. Genet. 104:388-398.

Martynov, V.V. & E.E. Khavkin. 2004 Two homologs of the *FLOWERING LOCUS C* gene from leaf mustard (*Brassica juncea*). Russian J. Plant Physiol. 51:234-240.

McNaughton, I.H. 1995. Swedes and rapes – *Brassica napus* (Cruciferae). In: Evolution of Crop Plants, 2nd Ed. J. Smartt and N.W. Simmonds (eds.). Longman Scientific & Technical, London, U.K. pp. 68-75.

Nei, M. & W.H. Li. 1979. Mathematical models for studying genetic variation in terms of restriction endonucleases. Proc. Natl. Acad. Sci. (USA) 76:5269-5273.

Osborn, T.C. & L. Lukens. 2003. The molecular genetic basis of flowering time variation in *Brassica* species. In: Biotechnology in Agriculture and Forestry, Vol. 52: Brassicas and Legumes from Genome Structure to Breeding. T. Nagata and S. Tabata (eds.). Springer-Verlag, Berlin, Germany. pp. 69-86.

Padmaja, K.L., N. Arumugam, V. Gupta, A. Mukhopadhyay, Y.S. Sodhi, D. Pental & A.K. Pradhan. 2005. Mapping and tagging of seed coat color and the identification of microsatellite markers for marker-assisted manipulation of the trait in *Brassica juncea*. Theor. Appl. Genet. 111:8-14.

Peakall, R. & P.E. Smouse. 2006. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288-295.

Plieske, J. & D. Struss. 2001. Microsatellite markers for genome analysis in Brassica. I. Development in *Brassica napus* and abundance in Brassicaceae species. Theor. Appl. Genet. 102:689-694.

Pradhan, A.K., S. Prakash, A. Mukhopadhyay & D. Pental. 1992. Phylogeny of *Brassica* and allied genera based on variation in chloroplast and mitochondrial DNA patterns: Molecular and taxonomic classifications are incongruous. Theor. Appl. Genet. 85:331-340.

Rajcan, I., K.J. Kasha, L.S. Kott & W.D. Beversdorf. 1999. Detection of molecular markers associated with linolenic and erucic acid levels in spring rapeseed (*Brassica rapa* L.). Euphytica 105:173-181.

Raybould, A.F., R.J. Mogg, R.T. Clarke, C.J. Gliddon & A.J. Gray. 1999. Variation and population structure at microsatellite and isozyme loci in wild cabbage (*Brassica oleracea* L.) in Dorset (UK). Genet. Res. Crop Evol. 46:351-360.

Riaz, A., G. Li, Z. Quresh, M.S. Swati & C.F. Quiros. 2001. Genetic diversity of oilseed *Brassica napus* inbred lines based on sequence-related amplified polymorphism and its relation to hybrid performance. Plant Breeding 120:411-415.

Rief, J.C., A.E. Melchinger & M. Frisch. 2005. Genetical and mathematical properties of similarity and dissimilarity coefficients applied in plant breeding and seed bank management. Crop Sci. 45:1-7.

Rohlf, F.J. 2005. NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, ver. 2.2. Exeter Publishing, Setauket, NY.

Salywon, A.M., M. Barber, N. Herling, W. Stewart & D.A. Dierig. 2004. Data mining for microsatellites in expressed sequence tags (ESTs) from *Arabidopsis thaliana* and *Brassica* species for use in *Lesquerella* (Brassicaceae). Abstracts of the Association for the Advancement of Industrial Crops Conference. September 19-22, 2004. Minneapolis, MN. p.20.

Sauer, J.D. 1993. Historical Geography of Crop Plants: A Select Roster. CRC Press, Boca Raton, FL.

Simpson, G.G., A.R. Gendall, & C. Dean. 1999. When to switch to flowering. Annu. Rev. Cell Dev. Biol. 99:519-550.

Skinner, D.Z., T. Loughin & D.E. Obert. 2000. Segregation and conditional probability association of molecular markers with traits in autotetraploid alfalfa. Molecular Breeding 6:295-306.

Snowdon, R.J. & W. Friedt. 2004. Molecular markers in *Brassica* oilseed breeding: Current status and future possibilities. Plant Breeding 123:1-8.

Song, K. & T.C. Osborn. 1992. Polyphyletic origins of *Brassica napus*: New evidence based on organelle and nuclear RFLP analysis. Genome 35:992-1001.

Sovero, M. 1993. Rapeseed, a new oilseed crop for the United States. In: New Crops. J. Janick and J.E. Simon (eds.). Wiley, New York, NY. pp. 302-307.

Thomas, P. 2003, Canola Grower's Manual. Canola Council of Canada. Winnipeg, Canada.

Tommasini, L., J. Batley, G.M. Arnold, R.J. Cooke, P. Donini, D. Lee, J.R. Law, C. Lowe, C. Moule, M. Trick & K.J. Edwards. 2003. The development of simple sequence repeats (SSR) markers to complement distinctiveness, uniformity and stability testing of rape (*Brassica napus* L.) varieties. Theor. Appl. Genet. 106:1091-1101.

Tonguç, M. & P.D. Griffiths. 2004. Genetic relationships of *Brassica* vegetables determined using database derived simple sequence repeats. Euphytica 137:193-201.

Uzunova, M.I. & W. Ecke. 1999. Abundance, polymorphism and genetic mapping of microsatellites in oilseed rape (*Brassica napus* L.). Plant Breeding 118:323-326.

Westman, A. & S. Kresovich. 1999. Simple sequence repeat (SSR)-based marker variation in *Brassica nigra* genebank accessions and weed populations. Euphytica 109:85-92.

Yu, C.Y., S.W. Hu, H.X. Zhan & A.G. Guo. 2005. Genetic distance revealed by morphological characters, isozymes, proteins and RAPD markers and their relationships with hybrid performance in oilseed rape (*Brassica napus* L.). Theor. Appl. Genet. 110:511-515.

Table 1. List of accessions selected from the *B. napus* germplasm collection at NCRPIS.

Plot ID[a]	*Accession*	*Variety*	*Source*[b]	*DAP to 50% flower*[c]
13N	Ames 6100	Jupiter	Canada	-
16N	Ames 15650	Arco C10-2	Netherlands	-
19N	Ames 15654	Bienvenu	United States	-
21	Ames 15939	Comet	Sweden	48
29N	Ames 19202	Krasnodarskii	Russia	-
31	Ames 19204	Evvin	Russia	41
32	Ames 19205	Kovalevskij	Ukraine	46
36	Ames 26653	Westar	United States	38
62	PI 311727	Bronowski	Poland	53
88	PI 436555	Gan You no. 2	China	59
92N	PI 443015	Gry	Norway	-
98N	PI 458610	Wilhelmsburger	New Zealand	-
100	PI 458919	Brio	France	44
108	PI 458930	Oro	Canada	48
109N	PI 458935	Brink	Sweden	-
113	PI 458940	Chisaya natane	Japan	83
114	PI 458941	Norin 16	Japan	68
119	PI 458948	Gisora	Germany	60
126	PI 458955	Prota	Germany	41
138	PI 458971	Romeo	France	39
173	PI 469756	Colza	South Korea	41
174	PI 469757	Colza 18 Miroc	South Korea	78
175	PI 469758	Dae cho sen	South Korea	68
179	PI 469762	Dong Hae 2	South Korea	73
189	PI 469772	Dong Hae 16	South Korea	75
193	PI 469776	Dong Hae 21	South Korea	60
205	PI 469788	Fertodi	South Korea	66
214	PI 469797	France 9	France	53
239	PI 469822	Iwashiro-natane	South Korea	66
243	PI 469826	Janetzkis	South Korea	63
276	PI 469859	Kuju 25	South Korea	40
298	PI 469881	Kuju 58	South Korea	76
303N	PI 469886	Lenora	South Korea	-
311	PI 469894	Mali	South Korea	35
327	PI 469911	Mokpo 5	South Korea	65
340	PI 469924	Mokpo 21	South Korea	59
346	PI 469930	Mokpo 27	South Korea	81
349	PI 469933	Mokpo 30	South Korea	69
356	PI 469940	Murame nadame	South Korea	36
371	PI 469955	Norin #4	Japan	79
391	PI 469975	Norin 21	Japan	45

Table 1. (continued).

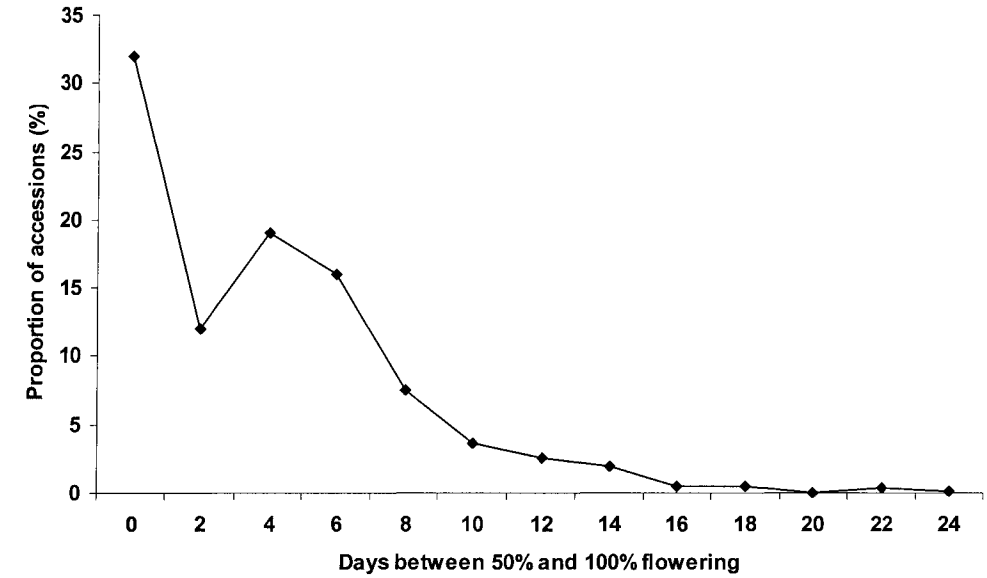
Plot ID[a]	*Accession*	*Variety*	*Source*[b]	*DAP to 50% flower*[c]
397	PI 469981	Norin 26	Japan	55
447	PI 470031	Su weon cheg	South Korea	79
457	PI 470041	Taiwan 2	Taiwan	40
489	PI 470075	7003-2B-38	South Korea	83
502	PI 478340	O 84	China	42
539N	PI 535866	Silesia	Czechoslovakia	-
573N	Ames 22547	Bolko	Poland	-
555	PI 542984	Tri-Bridger	United States	59
574	Ames 22548	Bronowski	Poland	51

[a]Field plot number, 'N' suffix denotes non-flowering , [b]Country origin in ARS-GRIN; Bienvenu and Westar's origin were changed to France and Canada, respectively (Diers and Osborn, 1994) in subsequent analyses, [c]Mean DAP values

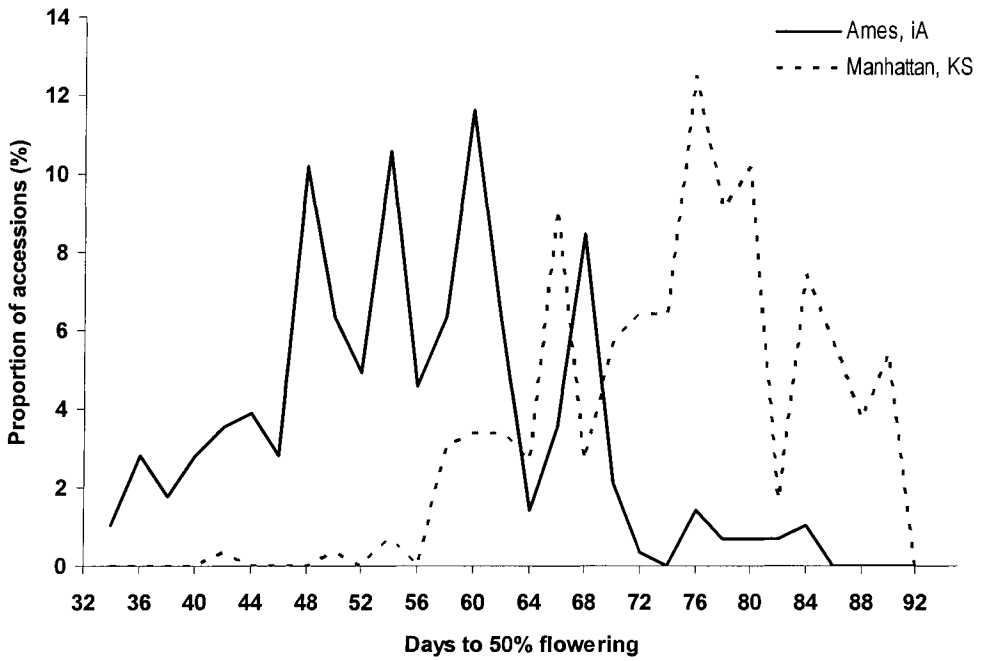
Table 2. List of the microsatellites analyzed, repeat motifs, size range of the amplified bands approximated from molecular weight markers, and the number of bands observed in the bulked samples. Sequences were obtained from BBSRC BrassicaDB, except as noted.

Microsatellite	Repeat	Size range (~bp)	Total no. bands
Na10-B08	$(CT)_{38}$	100-175	5
Na10-D03	$(GT)_{11}$	150-190	2
Na10-E02	$(GA)_{24}$	125-200	3
Na12-A02	$(CT)_{16}$	150-200	3
Na12-A07	$(GT)_{11}$	150-175	2
Na12-A08	$(GA)_{28}$	150-300	4
Na12-C07	$(CT)_{33}$	200-250	3
Na12-C08	$(CT)_{50}$	275-350	3
Na12-F03	$(GA)_{35}$	300-350	2
Na14-C12	$(AG)_{17}$	190-200	3
Na14-D07	$(CCG)_3$	150-175	2
Ni4-D09	$(CT)_{25}$	175-200	2
Ol10-A05	$(GA)_{43}$	110-275	7
Ol10-F11	$(GGC)_7$	150-175	2
Ol10-F12	$(CT)_{64}$	100-225	4
Ol11-C02	$(GT)_{11}$	140	1
Ol11-H02	$(AAC)_{18}$	180-200	2
Ol12-E03	$(CCG)_9$	110-250	3
Ol12-G04	$(TC)_{24}$	100-175	2
Ra2-D04	$(CA)_{14}$	160-190	2
Ra2-E03	$(CT)_{18}$	225-275	2
Ra2-E07	$(GA)_{19}$	100-170	4
Ra2-F11	$(CT)_{34}$	190-300	4
Ra2-G09	$(CT)_{19}$	200-300	3
Ra3-H10	$(GA)_{23}$	100-150	3
MR176*	n.d.	120-290	8
MR181*	$(AG)_{36}$	100-190	5
35D**	$(GA)_{13}$	200-250	2
59A1**	$(CA)_{11}$	450	1
25A**	$(CT)_{10}$	125-350	3
EST1***	n.d.	90	1
EST2***	n.d.	280	1
EST3***	n.d.	90-115	4

* Uzunova and Ecke (1999)
** Westman and Kresovich (1999)
*** A. Salywon (USDA-ARS, U.S. Water Conservation Laboratory, Phoenix, AZ)
n.d.-not determined



a



b

Figure 1. Graph of flowering time showing the proportion of 224 accessions versus (a) the number of days between 50% and 100% flowering in IA and (b) the number of days after planting to 50% flowering in IA and KS.

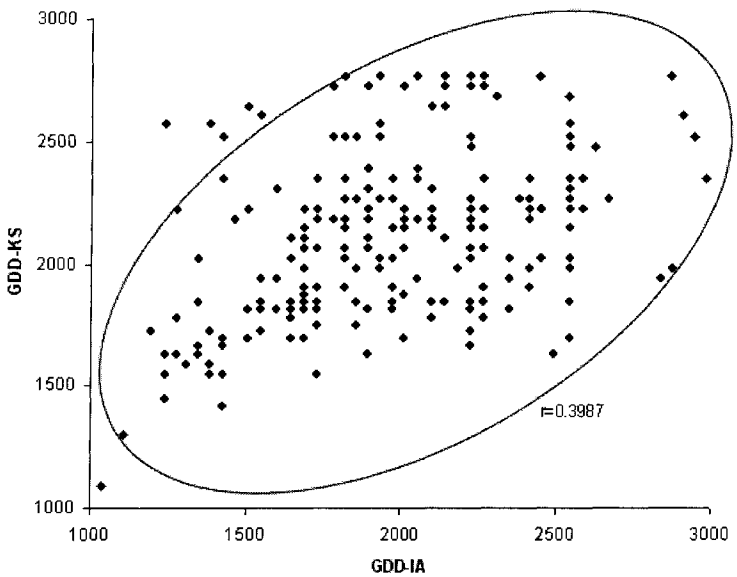


Figure 2. Scatterplot matrix of accumulated growing degree days (GDDs) to 50% flowering in Iowa and Kansas locations (density ellipse shown with α= 0.95).

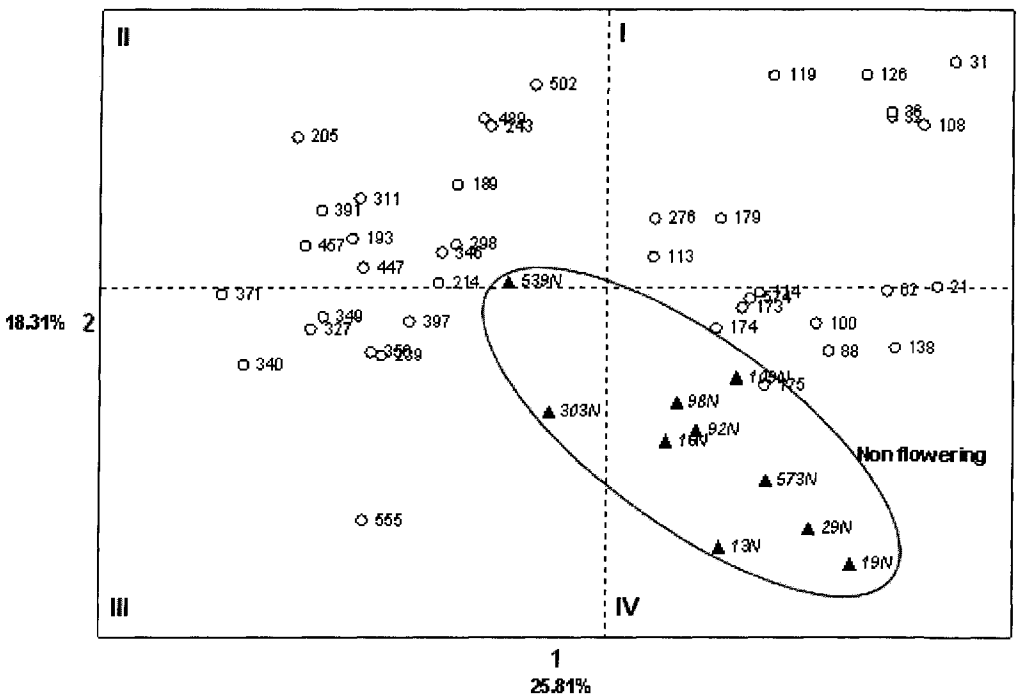


Figure 3. Plot of variation in flowering (o) and non-flowering (▲) accessions using eigenvectors derived from genetic distances by using shared alleles. Labels correspond to identification listed in Table 1.

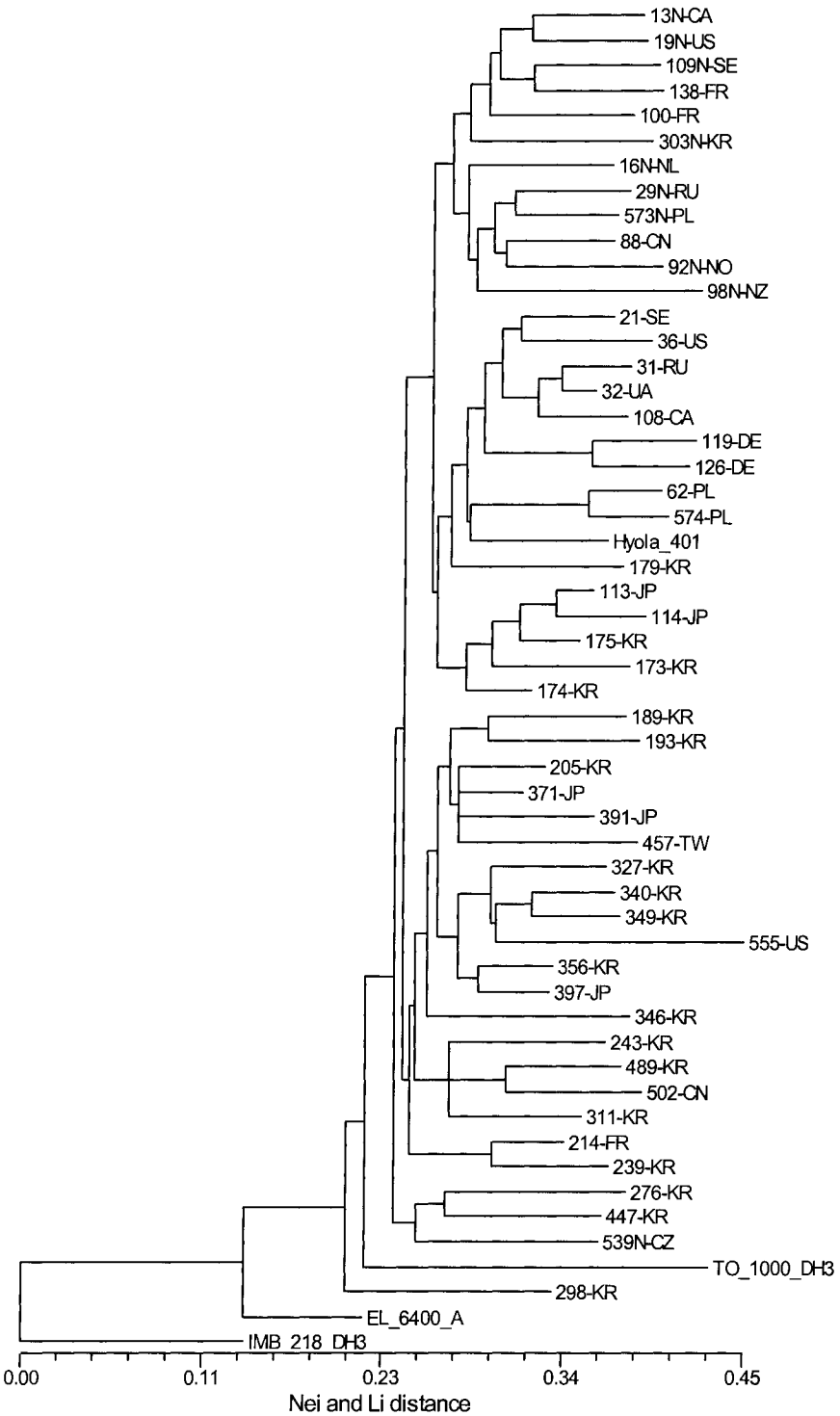


Figure 4. Neighbor-joining tree showing the relationships among the selected *B. napus* accessions and rapid cycling lines (*B. oleracea* TO 1000 DH3, *B. rapa* IMB 218 DH3, *B. napus* EL 6400 A). Labels correspond to plot numbers listed in Table 1 with added suffixes indicating country origin (in two letter ISO codes).

Table 3. Results of significant Mann-Whitney-U tests for association between non-specific and specific bands in five SSRs and flowering time.

SSR	Band no.	Size (~bp)	Band presence (%)		Prob.
			Non-flowering	Flowering	
EST2	1	280	70	47.5	0.01
Na10-B08	1	100	30	42.5	0.00
Ol11-C02	1	140	60	85	0.04
Ra2-E07	2	150	-	20	0.03
Ra2-E07	4	100	-	37.5	0.04
Ra2-F11	1	190	30	35	0.01

Table 4. AMOVA by life form (non-flowering vs. flowering) and by geographic regions (Europe, North America and Asia-Pacific-Oceania).

Source	df	SS	MS	Est. Var.	%	Φpt	Prob.
Life forms							
Among life forms	1	33.320	33.320	1.390	11%	0.111395	0.001
Within life forms	48	532.100	11.085	11.085	89%		
Geographical regions							
All accessions							
Among regions	2	40.207	20.103	0.677	6%	0.057157	0.002
Within regions	47	525.213	11.175	11.175	94%		
Flowering accessions only							
Among regions	2	46.755	23.378	1.498	12%	0.124442	0.001
Within regions	37	389.845	10.536	10.536	88%		
Non-flowering accessions only							
Among regions	2	22.000	11.000	0.179	2%	0.016722	0.399
Within regions	7	73.500	10.500	10.500	98%		

CHAPTER 3. SURVEY OF SEQUENCE VARIATION IN THE *FLOWERING LOCUS C* GENES (*FLC1* AND *FLC3*) IN ANNUAL AND BIENNIAL ACCESSIONS OF *BRASSICA NAPUS* L.

A paper to be submitted for publication in Genetics

Von Mark V. Cruz[1], J. Chris Pires[2], Joshua A. Udall[3], Randy C. Shoemaker[1, 4], E. Charles Brummer[1] and Candice A.C. Gardner[1, 5*]

Abstract

FLOWERING LOCUS C (*FLC*) is believed to play a central role in the autonomous control of flowering time in the model species *Arabidopsis*. Several copies of *FLC* have been identified in *Brassica*. Here we report the nucleotide diversity patterns in two copies of this gene in annual and biennial life forms of rapeseed (*Brassica napus* L.). Fifty accessions comprised of ten biennial and

---

[1]Department of Agronomy, Iowa State University, Ames, IA 50011

[2]Department of Biological Sciences, University of Missouri-Columbia, Columbia, MO 65211

[3]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011

[4]USDA-ARS Corn Insect and Crop Genetics Research Unit, Ames, IA 50011

[5]USDA-ARS Plant Introduction Research Unit and North Central Regional Plant Introduction Station, Ames, IA 50011

* *author for correspondence: email: gardnerc@iastate.edu*

forty annual types from three geographical regions were sampled from the germplasm collection of the North Central Plant Introduction Station (NCRPIS). It was observed that the variation in exon 4 to exon 6 of *FLC1* and *FLC3* as well as the number of identified haplotypes were lower in biennial types than in annuals. Considering the variation in exon 2 to 7 of *FLC3*, lowest variation was found near exons 4 and 5. Tests of neutrality using Tajima's *D* suggest the departure from neutral expectations at this gene region. Phylogenetic and association analyses of *FLC1* and *FLC3* haplotypes showed no distinct grouping of the representative accessions based on life forms or geographic regions.

Key words: canola, *FLC*, flowering time, life forms, MADS-box, nucleotide variation, rapeseed

Introduction

The mechanism of flowering has been the interest of many studies in the model species *Arabidopsis* and in horticultural and agronomic crops, such as almond, rice, wheat, barley, maize, soybean, and rapeseed. Understanding how flowering occurs may have significant impact in crop production by being able to produce varieties that flower early, late, or do not flower at all and possibly ones that could defy photoperiodism and vernalization requirements (Li et al., 2005; Silva et al., 2005).

Research on *Arabidopsis* revealed at least 54 loci that affect flowering (Levy and Dean, 1998). Although many of these genes have been sequenced, molecular studies to determine sequence variation in natural populations have been few. Studies that determined if there is an association between the nucleotide variation and flowering time include those in *FRIGIDA* (*FRI*) (Corre et al., 2002; Gazzani et al., 2003; Johanson et al., 2000) and *CRYPTOCHROME-2* (*CRY2*) (El-Assal et al., 2001). Within *Brassica*, the only pertinent survey has been of the *CONSTANS (CO)* and *CONSTANS-LIKE1* (*COL1)* genes in *Brassica nigra* L. (Koch) (Lagercrantz et al., 2002; Österberg et al., 2002).

In *B. rapa* L. and *B. napus,* where annual and biennial forms exist, the major QTL affecting flowering time was determined to correspond to the *Arabidopsis FLOWERING LOCUS C (FLC)* (Osborn et al., 1997). *FLC* is the key gene in floral initiation within the autonomous pathway of flowering-time control, and it encodes a MADS-box transcription factor that represses flowering (Kole et al., 2001; Sheldon et al., 2000; Simpson et al., 1999). More than 60% of the observed flowering time variation in *B. napus* was explained by QTL regions with mapped *FLC* copies within their confidence intervals (Udall et al., in press). In the cultivated species of *Brassica,* three copies of *FLC* are expected in the diploids (*B. oleracea* L., *B. nigra, B. rapa*) and up to eight in the amphidiploids (*B. carinata* Braun, *B. juncea* (L.) Czern., *B. napus*) (Schrantz et al., 2002; Tadege et al., 2001). One *FLC* homologous gene has been reported in *B. rapa* ssp. *pekinensis* (Lour.) Hanelt (Li et al., 2005), two homologs have been identified in *B. oleracea* (Lin et al., 2005) and in *B. juncea* (Martynov and Khavkin, 2004), and eight copies in a resynthesized *B. napus* (Udall, 2003). The several copies of *FLC* in *B. napus* were found to have an effect on increasing the variation in flowering time in a dosage-dependent manner (Osborn, 2004; Osborn and Lukens, 2003). Observations in resynthesized lines of *B. napus* also indicated that early and late-flowering lines differ in *FLC* structure and expression levels (Pires et al., 2004).

The sizes of *Brassica FLC,* from exons 2 to 7, range between 1000 to 2000 bp. Only exons 4 and 7 were reportedly variable in nucleotide length among the seven coding regions of the gene (Martynov and Khavkin, 2004; Schrantz et al., 2002). A recent survey of interspecific RFLP polymorphism in *FLC* in the cultivated *Brassica* indicated that the variation can successfully distinguish among the species (Martynov and Khavkin, 2005b). In this study, we wanted to determine if there are life-form specific *FLC* haplotypes. We tested the hypothesis that variation in two copies of *FLC* are associated with life forms of the crop using intraspecific sequence variation observed in germplasm accessions of *B. napus* representative of different life forms and flowering times.

Materials and Methods

Plant material and DNA extraction. Seeds of 40 annual and 10 biennial accessions of *B. napus* (Table 1) were sown in 10 cm square plastic pots containing Sunshine Mix no.1 (Sun Gro Horticulture, Bellevue, WA). Leaf tissues were harvested when the seedlings grown in a greenhouse were at the 3-4 leaf stage. Twelve to fifteen plants were sampled per accession. The leaves were lyophilized by freeze-drying and then stored in a -80°C freezer until use. Genomic DNA extraction was performed by using 1 g bulked tissue obtained from equal weights of freeze-dried leaf samples from individual plants within an accession. The bulked tissues were placed in 50 ml screw-cap polypropylene tubes, each containing approximately 1.5 ml glass beads. The tissue samples were ground to a powder by agitating the tubes on a paint shaker for 90 seconds. DNA was isolated by using the cetyltrimethyl ammonium bromide (CTAB) method. The extracted DNA was resuspended in 1X TE (1 mM Tris-HCl, 0.1 mM EDTA, pH 8.0) with 3 μl of 10 mg/ml RNAseA then incubated at 30°C for 30 minutes and placed on a LabQuake® shaker (Barnstead Intl., Dubuque, IA) at 4°C overnight. The samples were stored in -20°C until use.

PCR and DNA sequencing. Two different lengths of *Brassica FLC* were amplified with primers (Integrated DNA Technologies, Coralville, IA) from Schrantz et al. (2002) and Pires et al. (2004). For all 50 accessions, amplifications of the short fragments (exon 4 to exon 6) of *FLC1* and *FLC3* were done in 96-well Microseal™ polypropylene microplates (Bio-Rad Lab. Inc., Hercules, CA) with each sample well containing: 1 μl of genomic DNA (50 ng/μl), 8 μl of sterile ddH2O, 1 μl 10X PCR buffer (200 mM Tris-HCl, pH 8.4, 500 mM KCl), 0.5 μl dNTPs (2 mM), 0.3 μl MgCl (50 mM), 0.1 μl of each primer (50 μM), and 0.05 μl *Taq* polymerase (5 U/μl) (Invitrogen Corp., Carslbad, CA) on DNA Engine® (PTC-200™) thermal cyclers (Bio-Rad Lab. Inc.) with the following cycling profiles: 94°C for 4 min, then followed by 25 cycles of amplification at 94°C for 30 sec, 60°C for 45 sec, 72°C for 1

min, and final extension at 72°C for 10 min (Pires et al., 2004). A negative control containing only the PCR mix without the DNA sample was included in each microplate as well as three samples from rapid cycling *Brassica* (*B. oleracea* TO 1000 DH3, *B. rapa* IMB 218 DH3, *B. napus* EL 6400 A). For seven *B. napus* accessions, large fragments (exon 2 to exon 7) of *FLC3* were amplified using a forward primer with an additional four bases (CACC) at the 5' end, designed to allow directional cloning. The following cycling profile was followed to amplify the large fragments: 94°C for 4 min, then followed by 30 cycles of amplification at 94°C for 1 min 30 sec, 66°C for 1 min, 72°C for 1 min 30 sec, and final extension at 72°C for 10 min. All PCR products were separated in a 0.8% low melt agarose gel prepared in 1X TAE (40 mM Tris-acetate, 1 mM EDTA, pH 8.3) with incorporated ethidium bromide (0.46 μg/ml). Visualization of amplified products was done with a UV light box (254 nm wavelength) and photographed on a digital gel-documentation system. PCR bands were cut out from the gel and DNA extracted using QIAquick® Gel Extraction Kit (QIAGEN Inc., Valencia, CA). The purified PCR products from exon 4 to 6 were directly sequenced in 50 accessions.

Large fragment amplification products of *FLC3* were cloned for seven accessions to determine overall polymorphism across a larger region. The purified PCR product was cloned to a pENTR™/SD/D-TOPO® vector (Invitrogen Corp.). Chemical transformation was performed by incubating One Shot® TOP10 *E. coli* strains (Invitrogen Corp.) in ice for 30 minutes, then briefly transferring the cells to 42°C for 30 seconds. After transformation, the cells were incubated at 37 °C at 250 rpm for 1 hour. For each transformation reaction, 25 μl and 75 μl were spread on LB plates containing kanamycin (10 mg/ml). Eight colonies per accession were selected and transferred on a new LB plate after overnight incubation at 37°C. Transformants were analyzed by PCR and by restriction analysis using *NotI* enzyme (New England Biolabs, Ipswich, MA). Minipreps were performed using PureLink™ Quick Plasmid Miniprep Kit (Invitrogen Corp.).

All sequencing reactions were performed at the DNA Facility of the ISU Office of Biotechnology (http://www.dna.iastate.edu) using an ABI 3730 DNA Analyzer (Applied Biosystems Inc., Foster City, CA).

Sequence assembly and analysis: The consensus sequence for each *FLC* was obtained by assembling DNA fragments from several sequencing reactions (2-6 per accession) using Sequencher™ v.4.2.2 (Gene Codes Corp., Ann Arbor, MI). Multiple sequence alignments were done using ClustalW implemented in BioEdit v.7.0.5.2 (Hall, 1999) with manual adjustments. Location of exons and introns was determined by aligning the sequences obtained to published *FLC* sequences in *B. rapa* (AY115678, AY115677) and *B. oleracea* (AY115674, AY115673). One accession (ID175) without full length sequence of *FLC1* exon 4 to 6 was removed in subsequent analyses. The determination of nucleotide diversity, tests of neutrality, linkage disequilibrium, and genetic differentiation were conducted using DnaSP v.4.10.6 (Rozas et al., 2003). Likewise, the software was used to perform sliding window analyses to determine regions in the gene where genetic differentiation was appeared to be enhanced compared to neutral expectations. The sliding window analyses of the *FLC* sequences were conducted using a fixed window size of 50 bp and a step size of 15 bp. The rate of synonymous (*Ks*) and non-synonymous (*Ka*) nucleotide substitutions was computed. Also, the nucleotide polymorphisms at all sites were determined by estimating polymorphism index (*Pi*), the number of segregating sites (*S*), number of singletons (variants that appear only once in the total sample), the number of haplotypes from the consensus sequences (*H*), and haplotype diversity (*Hd*). Tajima's *D* statistic (Tajima, 1989) and Fu and Li's *F* statistic (Fu and Li, 1993) were also computed to determine if the intraspecific patterns of variation detected in the *FLC* region analyzed were in agreement with neutral predictions. Tajima's *D* statistic is based on the standardized difference between two estimates of the neutral parameter ($\theta = 4N_e\mu$), one based on the observed average number of pairwise differences (k) and the other on the number of segregating sites in the sample (*S*). Fu and Li's (1993)

F statistic is based on the standardized difference between k and the number of external mutations. Coalescent simulations with 10000 replicates were performed to determine the expected distribution and significance of Tajima's *D* and Fu and Li's *F*. Tests of genetic differentiation between groups of life forms and among the designated geographical regions of origin were performed by computing the average heterozygosity among groups (*Hs*), the average number of differences between sequences from within groups (*Ks*), the coefficient of gene differentiation (*Gst*), and equilibrium fixation index (*Fst*). The statistical significance of *Hs* and *Ks* was tested by means of permutations tests with 10000 replicates. In the short fragments, because of the small region analyzed, statistical estimates were reported considering both exon and intron sequences. In all sequences, *Ks* and *Ka* values were derived using exons only.

Relationships among the identified haplotypes were examined by computing the genetic distances among them using PHYLIP v.3.6 (Felsenstein, 2005). Clustering of the haplotypes was performed using BIO-NJ neighbor-joining (Gascuel, 1997), an algorithm which improves over the neighbor-joining algorithm (NJ) of Saitou and Nei (1987) when using nucleotide sequences. Visualization of resulting trees was done using the software TreeView v.1.6.6 (Page, 1996). Case-control association tests were performed to answer the null hypothesis that allele frequencies at the candidate locus do not depend on phenotype (i.e. flowering or non-flowering). To test for association between the nucleotide sites and the phenotypes, the structure of the population and the ancestry of the individuals in the sample were initially estimated under a Bayesian framework implemented by the software Structure v.2.0 (Pritchard et al., 2000a; Pritchard et al., 2000b). A 50000 iteration burn-in period and $5x10^5$ Monte Carlo Markov Chain iterations were used as final settings for data gathering in Structure. The estimates from Structure were then used in the software STRAT v.1.1 (Pritchard et al., 2000b) to determine presence of significant associations to the phenotypes at each nucleotide site in the haplotype sequences.

Results

Successful amplification was obtained in all accessions with primers for *FLC1* and *FLC3*. The large cloned fragments of *FLC3* in seven accessions were more than 1.5kb and were trimmed to a common 1553bp sequence corresponding to nucleotides in exon 2 to 7. In the more extensive survey of 50 accessions, the short regions amplified had an average of 478bp sequence in *FLC1* and 532bp in *FLC3*. After alignment, the short fragments of *FLC1* and *FLC3,* corresponding to nucleotides in exon 4 to exon 6, were trimmed to 403bp and 441bp overlapping regions, respectively and analyzed.

Results indicated fully conserved exon sequences across annual and biennial types. The organization of the variation in the long and short sequence fraction of *FLC* from the different life forms and three populations are presented below and summarized in Tables 1 and 2.

Variation in long fragments of FLC3. The larger cloned fragments of *FLC3* have an average size of 1.5kb and contain exon 2 to exon 7 and introns 2 to 6. Considering all sampled sequences, there was a total of 115 singletons and 193 polymorphic sites observed (Table 2). The computed rate of synonymous substitutions (*Ks*) is higher than the nonsynonymous substitutions (*Ka*) in both annual and biennial types. The total nucleotide diversity, *Pi*, is 0.049. Because of fewer accessions representative of the larger *FLC3* fragment, only grouping on life form was considered. There was higher nucleotide diversity in the annual accessions (Pi=0.058) than the biennials (Pi=0.019). The shorter fragment analyzed in 50 accessions and presented above was found to correspond to one of the regions where there is low nucleotide diversity overall. There is more variability in gene regions near exon 3 and exon 6 considering the representative accessions (Figure 1). Among the introns, low variability was observed in introns 3 and 5, while highest variation was observed in intron 2. Sliding window analysis of Tajima's *D* indicated that most of the sequenced regions have negative values of *D,* except for the regions surrounding exons 4 and 5, and near exon 7.

Nucleotide polymorphism in short fragments of FLC1. The estimates of nucleotide polymorphism between life forms indicate the higher variation in *FLC1* in annuals than the biennial types. A total of 28 and 5 singletons were observed in annual and biennial types, respectively. There was a total of 18 haplotypes identified in annual types and 7 in biennials. Among geographic regions, nucleotide diversity ranged from 0.006 (North America) to 0.012 (Europe) and number of haplotypes ranged from 4 (North America) to 13 (Asia-Pacific). In the designated groups (i.e. life forms and geographical regions), the *Ka* is higher than the *Ks* except in the North American group.

Nucleotide polymorphism in short fragments of FLC3. A similar trend to *FLC1* in nucleotide variation was observed in *FLC3*. The annual types have higher nucleotide diversity (*Pi*=0.060) than the biennials (*Pi*=0.023) as well as the number of singletons (36 in annuals vs. 34 in biennials) and number of haplotypes (28 in annuals vs. 9 in biennials). Among the regional groupings, the nucleotide diversity ranged from 0.025 (Europe) to 0.077 (Asia-Pacific) and number of haplotypes from 4 (North America) to 26 (Asia-Pacific). In contrast to what was observed in *FLC1,* the *Ks* values are higher than the *Ka* in all groupings.

The average number of nucleotide substitutions per site (*Dxy*) and number of net nucleotide substitutions per site (*Da*) between population are additional parameters that can be used to dissect interpopulation variation and divergence (Nei, 1987). Analyses of these parameters in annual and biennial groups indicate larger values of *Dxy* in intron 5 in both *FLC* copies (Figure 2). The computed *Dxy* and *Da* values between annual and biennial groups were greater in *FLC3* than in *FLC1* (highest *Dxy* in *FLC1* = 0.02 vs. highest *Dxy* in *FLC3* = 0.13 (Figures 2a and 2b), however the values obtained are very close to zero.

Between the *FLC* copies surveyed, a slightly higher overall polymorphism level was found in *FLC3.* There were more singletons and polymorphic sites observed in *FLC3* (33) than in *FLC1* (28)

as well as the total number of haplotypes (36 in *FLC3* vs. 23 in *FLC1*). The lists of haplotypes are shown in Figures 3 and 4. Phylogenetic analyses of the haplotypes observed in both gene copies did not reveal groupings by life form or assigned geographic origins (Figure 5). Most of the haplotypes were observed in both life forms. The haplotypes observed in biennials did not group in a single cluster as shown by the cladograms.

Tests of association between nucleotide sites and life form indicated that there are more significantly associated nucleotide sites in *FLC3* haplotypes than in *FLC1*. The results suggest that these seem to co-segregate or associate with annualism or biennialism. There were only three sites found significant in *FLC1* haplotypes while there are thirty significant sites in *FLC3* haplotypes. The significant sites in *FLC3* however, when used in phylogenetic analysis are still not sufficient to delimit groups of biennial life forms (tree not shown).

Neutrality tests and linkage disequilibrium. The computed Tajima's *D* values were negative for all groups in the two *FLC* copies (Table 4). Between copies, more negative values were observed in *FLC1*. Tajima's *D* values are significant in *FLC1* and in the annuals and Asia-Pacific groups of the gene. Results of Fu and Li's test corroborate significant values of the *D* statistic in *FLC1*'s aforementioned groups. The low values of LD (*Zns*) in both *FLC1* and *FLC3* suggest high recombination rates, and this is also evident in the numerous haplotypes obtained. Overall, LD values in *FLC3* are lower than *FLC1*.

Population differentiation between designated groupings. Observed values of *Fst* and *Gst* suggest very little genetic differentiation between life forms and among geographic regions on *FLC1* and *FLC3* (Table 5). Significant differences however were obtained on the estimates of *Hs* and *Ks* between Asia-Pacific and Europe, and between life forms using *FLC3*.

Discussion

Tests of neutrality on whole sequences of *FLC3* in selected accessions indicate that regions near exon 4 and 6 have positive *D* values suggesting traces of selection (Figure 1). We focused on this region on the 50 accessions of *B. napus* because it has been suggested that specific phenotypic differences between species (or in the current situation between life forms) are more likely to be observed in such genomic regions where selection has been suggested to occur (Nielsen, 2001). Furthermore, neutral sites near the direct target of selection may also be informative as these regions are also expected to have enhanced levels of genetic differentiation (Charlesworth et al., 1997).

Tests of neutrality on the short fragments (exon 4 to 6) of *FLC1* and *FLC3* corroborate the result observed on the long sequences (exon 2 to exon 7). In *FLC1*, Tajima's *D* values were mostly negative along the stretch of the gene in annual and biennial types as indicated by sliding window analysis (not shown). In contrast, regions near *FLC3* intron 4 in annuals and exon 5 in biennials have positive Tajima's *D*. These results suggest that some kind of purifying selection is acting in *FLC1* while balancing selection in *FLC3* (Innan and Stephan, 2000). This is also suggested by the higher *Ka* values relative to *Ks* in *FLC1* (Table 3). Overall, representative MADS-box regulatory genes (where *FLC* has been classified) were found to be evolving faster than structural genes based on relative non-synonymous to synonymous substitution rates (Barrier et al., 2001).

Results of this survey indicated similar levels of variability in regions of *FLC1* and *FLC3* in *B. napus* accessions. It was found that the gene region (exon 4 to 6) analyzed in the 50 selected accessions are relatively conserved, and have low polymorphism compared to other parts of the gene (Figure 1). Levels of polymorphism have previously been described in other *Brassica.* In particular, regions near exon 4 in the short region sequenced, have very little variability and have been proposed as critical in the functional position of the *FLC* gene in *B. rapa* and *B. juncea* (Schrantz et al., 2001;

Martynov and Khavkin, 2004). In the current survey, there was no variability observed in exon 5 across all accessions analyzed.

Intron variability in the copies of *FLC* analyzed indicated higher variation in intron 5 than intron 4. The low variability that was observed in intron 4 and in the exon regions corroborates the results of *FLC* polymorphism reported in *B. juncea* (Martynov and Khavkin, 2004). Intron variability was found sufficient to discriminate among different species of *Brassica* (Martynov and Khavkin, 2004) but not between intraspecific groups of life forms and geographic regions in this study.

Differentiation between life forms was not observed even when the levels of polymorphism in the *FLC3* long fragments were considered. The biennial accessions were not grouped together in phylogenetic trees derived from the long fragments (tree not shown). This result is similar to what was determined in recent RFLP studies on *FLC3* where cluster analysis using restriction patterns did not discern life forms from different latitudes (Martynov and Khavkin, 2005b). Similarly, the analysis of haplotypes that were identified from the consensus sequences in both *FLC1* and *FLC3* did not suggest any meaningful clusters with most biennial types found interspersed among annual types (Figure 5). Both of the life forms have shared haplotypes, in particular Hap 2 and Hap 4 in *FLC1,* and Hap 1 in *FLC3.* This study did not examine the complete *FLC* gene sequence or all of the *FLC* copies known in *B. napus* for variation. It would be interesting to determine if our findings would apply to all *FLC* copies. Also sequences upstream of *FLC* might be of interest. In the *B. nigra CO* gene, allelic variation associated with flowering time was suggested to reside outside the coding region of the gene (Österberg et al., 2002). It was subsequently determined to be 3.5 kb upstream of *B. nigra CO.* Recently, in maize, *cis*-acting elements controlling differences in phenotypes by influencing the expression of *teosinte branched1* (*tb1*) were determined to be very distant (> 41 kb) upstream of the gene (Clark et al., 2006).

The relative importance of *FLC3* over *FLC1* in influencing the diversity in flowering time is not addressed in this study. However, previous observations indicated that early and late-flowering *B.*

napus lines differ in *FLC3* structure and expression levels, which was in contrast with what was found in *FLC1* (Pires et al., 2004). In addition, it was determined that it is the parental copy of *FLC3* from *B. oleracea* that has higher expression levels and influences the delay in flowering more than the copy from *B. rapa.* It remains to be determined whether subfunctionalization of the parental *FLC* copies in *B. napus* occurs. Preliminary studies in another flowering-time gene, *CONSTANS,* in *B. napus* suggested that expression of the gene occurs not only in leaves but also in cotyledons and roots (Robert et al., 1998). In other allopolyploids such as cotton (*Gossypium* sp.), it was shown that different parental gene copies of alcohol dehydrogenase A (*adhA*) homoeologs are expressed in different tissues (Adams and Wendel, 2005). Between the *B. napus FLC* genes, no study of such type has yet been conducted comparing annual and biennial types.

Finally, the computed measures of genetic differentiation (*Gst, Fst*) suggest that, among regions and between life forms, there is low genetic differentiation. Significant values ($0.001<p<0.05$) were obtained only in *FLC3* on *Ks* between annual and biennials and *Ks* between Asia-Pacific and Europe. The significant number of differences between sequences in those groupings however did not translate to significant values of *Gst* or *Fst.* The values obtained for the coefficient of gene differentiation were comparable to that previously estimated using thirty microsatellite markers using the same set of annual and biennial accessions in this study (Cruz et al., accepted). *Brassica napus* is believed to be in existence for just 400 hundred years; this is fairly recent compared to other crops (Gómez-Campo and Prakash, 1999; McNaughton, 1995). Differentiation among *B. napus* types (i.e. fodder, spring, and winter) has been reportedly detected using microsatellites on elite and breeding lines (Tommasinni et al., 2003) and by restriction patterns and whole sequences of the *CONSTANS* and *COL* gene (Axelsson, 2000; Lagercrantz et al., 2002; Martinov and Khavkin, 2005a; Robert et al., 1998). However, this is not possible with the level of conservation in the *FLC* region analyzed and of the *FLC* copies studied herein.

Conclusions

We have determined that there is little polymorphism in the *FLC* gene copies in germplasm materials comprised of commercial cultivars and landraces. In particular, exon 5 was found to be highly conserved across accessions. It was determined that there is a slightly higher overall variation in *FLC3* than in *FLC1* considering the short gene region surveyed, which corresponds to about one-third of the length of exons 2 to 7. It was found out that there was no significant genetic differentiation between annual and biennial types of *B. napus*. Likewise, estimates of population differentiation also indicated little differences among the designated geographical regions from where the representative accessions are believed to have originated.

Acknowledgements

This is a journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 1018, and was supported by Hatch Act and the State of Iowa. We would like to thank Dr. T.C. Osborn (Seminis) for advice during the initial phases of the study, C. Grover, J. Hayes, L. Lincoln, J. Schlueter and S. Schlueter and the dedicated staff of Shoemaker's Lab. at the USDA-ARS Corn Insect and Crop Genetics Research Unit for technical assistance in the laboratory, and the staff of USDA-ARS NCRPIS for help in seed preparation, greenhouse planting and maintenance. We are also grateful to Dr. L. Lukens and Dr. K.R. Lamkey for reviewing the draft manuscript and the support from Fulbright program grant of the U.S. Department of State to VMVC. Mention of commercial brand names in this paper does not constitute an endorsement of any product by the U.S. Department of Agriculture or cooperating agencies.

References

Adams, K.L. and J.F. Wendel. 2005. Novel patterns of gene expression in polyploid plants. Trends in Genetics 21:539-543.

Axelsson, T. 2000. Evolution of polyploid *Brassica* genomes: Genome structure and the evolution of duplicated genes. Doctoral thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden.

Barrier, M., R.H. Robichaux, and M.D. Purugganan. 2001. Accelerated regulatory gene evolution in an adaptive radiation. Proc. Natl. Acad. Sci. (USA) 98:10208-10213.

Charlesworth, B., M. Nordborg, and D. Charlesworth. 1997. The effects of local selection, balanced polymorphism and background selection on equilibrium patterns of genetic diversity in subdivided populations. Genet. Res. 70:155-174.

Clark, R.M., T.N. Wagler, P. Quijada, and J. Doebley. 2006. A distant upstream enhancer at the maize domestication gene *tb1* has pleiotropic effects on plant and inflorescent architecture. Nature Genetics 38:594-597.

Corre, V.L., F. Roux, and X. Reboud. 2002. DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: Extensive nonsynonymous variation is consistent with local selection for flowering time. Mol. Biol. Evol. 19:1261-1271.

Cruz, V.M.V., R. Luhman, L.F. Marek, C.L. Rife, R,C. Shoemaker, E.C. Brummer, and C.A.C. Gardner. Characterization of flowering time and SSR marker analysis of spring and winter type *Brassica napus* L. germplasm (accepted for publication in Euphytica)

El-Assal, S.E., C. Alonso-Blanco, A.J.M. Peeters, V. Raz, and M. Koornneef. 2001. A QTL for flowering time in *Arabidopsis* reveals a novel allele for *CRY2*. Nature Genetics 29:435-440.

Felsenstein, J. 2005. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. Department of Genome Sciences, University of Washington, Seattle, WA.

Fu, Y.X. and W.H. Li. 1993. Statistical tests of neutrality of mutations. Genetics 133:693-709.

Gascuel O. 1997. BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. Mol. Biol. Evol. 14:685-695.

Gazzani, S., A.R. Gendall, C. Lister, and C. Dean. 2003. Analysis of the molecular basis of flowering time variation in *Arabidopsis* accessions. Plant Physiol. 132:1107-1114.

Gómez-Campo, C. and S. Prakash. 1999. Origin and domestication. In: Biology of *Brassica* Coenospecies. (Developments in Plant Genetics and Breeding, vol. 4). C. Gómez-Campo (ed.). Elsevier Science, Amsterdam, The Netherlands. pp. 33-58.

Hall, T.A. 1999. BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids Symp. Ser. 41:95-98.

Innan, H. and W. Stephan. 2000. The coalescent in an exponentially growing metapopulation and its application to *Arabidopsis thaliana.* Genetics 155:2015-2019.

Johanson, U., J. West, C. Lister, S. Michaels, R. Amasino, and C. Dean. 2000. Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290:344-347

Kole, C., P. Quijada, S.D. Michaels, R.M. Amasino, and T.C. Osborn. 2001. Evidence for homology of flowering-time genes *VFR2* from *Brassica rapa* and *FLC* from *Arabidopsis thaliana.* Theor. Appl. Genet. 102:425-430.

Lagercrantz, U., M.K. Österberg, and M. Lascoux. 2002. Sequence variation and haplotype structure at the putative flowering-time locus *COL1* of *Brassica nigra.* Mol. Biol. Evol. 19:1474-1482.

Levy, Y.Y. and C. Dean. 1998. Control of flowering time. Curr. Op. Plant Biol. 1:49-54.

Li, Z., L. Zhao, C. Cui, G. Kai, L. Zhang, X. Sun, and K. Tang. 2005. Molecular cloning and characterization of an anti-bolting related gene (*BrpFLC*) from *Brassica rapa* ssp. *pekinensis.* Plant Science 168:407-413.

Lin, S., J. Wang, S. Poon, C. Su, S. Wang, and T. Chiou. 2005. Differential regulation of *FLOWERING LOCUS C* expression by vernalization in cabbage and *Arabidopsis*. Plant Physiology 137:1037-1048.

Martynov, V.V. and E.E. Khavkin. 2004. Two homologs of the *FLOWERING LOCUS C* gene from leaf mustard (*Brassica juncea*). Russian J. Plant Physiol. 51:234-240.

Martynov, V.V. and E.E. Khavkin. 2005a. Polymorphism of the *CONSTANS* gene in *Brassica* plants. Russian J. Plant Physiol. 52:242-248

Martynov, V.V. and E.E. Khavkin. 2005b. RFLP analysis of the *FLOWERING LOCUS C* gene in six *Brassica* species. Russian J. Plant Physiol. 52:352-258.

McNaughton, I.H. 1995. Swedes and rapes – *Brassica napus* (Cruciferae). In: Evolution of crop plants. 2nd. Ed. J. Smartt and N.W. Simmonds (eds.). Longman Scientific & Technical, Essex, U.K. pp. 68-75.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nielsen, R. 2001. Statistical tests of selective neutrality in the age of genomics. Heredity 86:641-647.

Osborn, T.C. 2004. The contribution of polyploidy to variation in *Brassica* species. Physiologia Plantarum 121:531-536.

Osborn, T.C. and L. Lukens. 2003. The molecular genetic basis of flowering time variation in *Brassica* species. In: Biotechnology in Agriculture and Forestry, Vol. 52: Brassicas and Legumes from Genome Structure to Breeding. T. Nagata and S. Tabata (eds.). Springer-Verlag, Berlin, Germany. pp. 69-86.

Österberg, M.K. O. Shavorskaya, M. Lascoux, and U. Lagercrantz. 2002. Naturally occurring indel variation in the *Brassica nigra COL1* gene is associated with variation in flowering time. Genetics 161:299-306.

Page, R.D.M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12: 357-358.

Pires, J.C., J. Zhao, M.E. Schranz, E.J. Leon, P.A. Quijada, L.N. Lukens, and T.C. Osborn. 2004. Flowering time divergence and genomic rearrangements in resynthesized *Brassica* polyploids (Brassicaceae). Biol. J. Linn. Soc. 82:675-688.

Pritchard, J.K., M. Stephens, and P. Donnelly, 2000a. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Pritchard, J.K., M. Stephens, N.A. Rosenberg, and P. Donnelly. 2000b. Association mapping in structured populations. Am J. Hum. Genet. 67:170-181

Quijada, P.A., J.A. Udall, B. Lambert, and T.C. Osborn. in press. Identification of genomic regions from winter germplasm that affect seed yield and other complex traits in hybrid rapeseed (*Brassica napus* L.).

Robert, L.S., F. Robson, A. Sharpe, D. Lydiate, and G. Coupland. 1998. Conserved structure and function of the *Arabidopsis* flowering time gene *CONSTANS* in *Brassica napus*. Plant Mol. Biol. 37:763-772.

Rozas, J., J.C. Sánchez-DelBarrio, X. Messeguer, and R. Rozas. 2003. DnaSP, DNA polymorphism analyses by the coalescent and other methods. Bioinformatics 19:2496-2497.

Saitou N. and M. Nei. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406-25.

Schrantz, M.E., P. Quijada, S. Sung, L. Lukens, R. Amasino, and T.C. Osborn. 2002. Characterization of the replicated flowering time gene *FLC* in *Brassica rapa*. Genetics 162:1457-1468.

Silva, C., J. Garcia-Mas, A.M. Sanchez, P. Arús, and M.M. Oliveira. 2005. Looking into flowering time in almond (*Prunus dulcis* (Mill) D.A. Webb): The candidate gene approach. Theor. Appl. Genet. 110:959-968.

Simpson, G.G., A.R. Gendall, and C. Dean. 1999. When to switch to flowering. Annu. Rev. Cell. Dev. Biol. 99:519-550.

Sheldon, C.C., D.T. Rouse, E.J. Finnegan, W.J. Peacock, and E.S. Dennis. 2000. The molecular basis of vernalization: The central role of *FLOWERING LOCUS C* (*FLC*). Proc. Nat. Acad. Sci. (USA) 97:3753-3758.

Tadege, M., C.C. Sheldon, C.A. Helliwell, P. Stoutjesdijk, E.S. Dennis, and W.J. Peacock. 2001. Control of flowering time by FLC orthologues in *Brassica napus*. The Plant Journal 28:545-553.

Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123:585-595.

Tommasini, L., J. Batley, G.M. Arnold, R.J. Cooke, P. Donini, D. Lee, J.R. Law, C. Lowe, C. Moule, M. Trick, and K.J. Edwards, 2003. The development of simple sequence repeats (SSR) markers to complement distinctiveness, uniformity and stability testing of rape (*Brassica napus* L.) varieties. Theor. Appl. Genet. 106:1091-1101.

Udall, J.A. 2003. A genetic study of oilseed *Brassica napus*: Mapping chromosome rearrangements and quantitative trait loci. Ph.D. dissertation. University of Wisconsin, Madison, WI.

Udall, J.A., P.A. Quijada, B. Lambert, and T.C. Osborn. in press. Identification of alleles from unadapted germplasm affecting seed yield and other quantitative traits in hybrid oilseed *Brassica napus* L.

Table 1. List of NCRPIS *B. napus* accessions analyzed.

Accession[a]	*Variety*	*Source*[b]
Ames 6100*	Jupiter	Canada
Ames 15650*	Arco C10-2	Netherlands
Ames 15654*	Bienvenu	United States
Ames 15939	Comet	Sweden
Ames 19202*	Krasnodarskii	Russia
Ames 19204	Evvin	Russia
Ames 19205	Kovalevskij	Ukraine
Ames 22547*	Bolko	Poland
Ames 22548	Bronowski	Poland
Ames 26653	Westar	United States
PI 311727	Bronowski	Poland
PI 436555	Gan You no. 2	China
PI 443015*	Gry	Norway
PI 458610*	Wilhelmsburger	New Zealand
PI 458919	Brio	France
PI 458930	Oro	Canada
PI 458935*	Brink	Sweden
PI 458940	Chisaya natane	Japan
PI 458941	Norin 16	Japan
PI 458948	Gisora	Germany
PI 458955	Prota	Germany
PI 458971	Romeo	France
PI 469756	Colza	South Korea
PI 469757	Colza 18 Miroc	South Korea
PI 469758	Dae cho sen	South Korea
PI 469762	Dong Hae 2	South Korea
PI 469772	Dong Hae 16	South Korea
PI 469776	Dong Hae 21	South Korea
PI 469788	Fertodi	South Korea
PI 469797	France 9	France
PI 469822	Iwashiro-natane	South Korea
PI 469826	Janetzkis	South Korea
PI 469859	Kuju 25	South Korea
PI 469881	Kuju 58	South Korea
PI 469886*	Lenora	South Korea
PI 469894	Mali	South Korea
PI 469911	Mokpo 5	South Korea
PI 469924	Mokpo 21	South Korea
PI 469930	Mokpo 27	South Korea
PI 469933	Mokpo 30	South Korea
PI 469940	Murame nadame	South Korea
PI 469955	Norin #4	Japan
PI 469975	Norin 21	Japan
PI 469981	Norin 26	Japan
PI 470031	Su weon cheg	South Korea
PI 470041	Taiwan 2	Taiwan
PI 470075	7003-2B-38	South Korea
PI 478340	O 84	China
PI 535866*	Silesia	Czechoslovakia
PI 542984	Tri-Bridger	United States

[a]Asterisks denote biennial type, [b]Source as listed in USDA-ARS GRIN database

Table 2. Estimates of nucleotide variation at *FLC3* exon 2 to 7

Group	Sites	*Ks*	*Ka*	*S*	Singletons	*Pi*	θw, sd	*H*	*Hd*, sd
Annuals	1502	0.118	0.082	175	29	0.058	0.056, 0.028	5	1.000, 0.126
Biennials	1505	0.051	0.024	29	102	0.019	0.019, 0.014	2	1.000, 0.500
All	1499	0.097	0.068	193	115	0.049	0.053, 0.024	7	1.000, 0.076

Ka= rate of synonymous substitutions; *Ks*= rate of non-synonymous substitutions; *S*= no. segregating (polymorphic) sites; *Pi*= polymorphism index; θw –theta computed from *S*; sd= standard deviation - without recombination); *H*= no. haplotypes; *Hd*= haplotype diversity;

Table 3. Estimates of nucleotide variation at *FLC1* and *FLC3* exon 4 to 6.

Group	Sites	*Ks*	*Ka*	*S*	Singletons	*Pi*	θw, sd	*H*	*Hd*, sd
FLC1									
Life form									
Annuals	380	0.011	0.017	38	28	0.009	0.023, 0.008	18	0.822, 0.058
Biennials	384	0.000	0.003	15	5	0.013	0.014, 0.006	7	0.933, 0.062
Regions									
AP	381	0.006	0.015	37	28	0.010	0.024, 0.008	13	0.833, 0.063
NA	385	0.030	0.000	5	5	0.006	0.006, 0.003	4	0.900, 0.161
EU	382	0.009	0.019	25	16	0.012	0.019, 0.007	12	0.917, 0.064
All	379	0.009	0.015	45	28	0.010	0.026, 0.008	23	0.856, 0.046
FLC3									
Life form									
Annuals	388	0.082	0.054	116	36	0.061	0.070, 0.021	29	0.946, 0.027
Biennials	410	0.064	0.031	45	34	0.028	0.039, 0.016	9	0.978, 0.054
Regions									
AP	392	0.110	0.062	124	36	0.077	0.080, 0.026	26	0.998, 0.010
NA	407	0.075	0.041	60	58	0.059	0.070, 0.036	4	0.900, 0.161
EU	401	0.035	0.026	55	37	0.025	0.041, 0.015	11	0.875, 0.081
All	386	0.084	0.050	115	33	0.054	0.067, 0.019	36	0.947, 0.024

Regions: AP= Asia-Pacific; NA= North America; EU= Europe. See Table 2 for other abbreviations

Table 4. Tests of neutrality and measure of linkage disequilibrium in different groupings.

Group	D_{Tajima}	$F_{\text{Fu\&Li}}$	Z_{nS}
FLC1			
Annuals	-2.393**	-4.164**	0.153
Biennials	-0.682	0.074	0.354
AP	-2.337**	-3.548**	0.178
EU	-1.824*	-2.179	0.134
NA	-1.384***	-1.236	0.375
All	-2.348**	-3.886**	0.103
FLC3			
Annuals	-1.423	-1.546	0.078
Biennials	-1.525	-1.830	0.042
AP	-1.269	-1.427	0.078
EU	-1.866*	-2.237	0.227
NA	-1.208	-1.310	0.890
All	-1.596	-1.920	0.073
***FLC3* long fragment**	-0.705	-0.574	0.041

Significance: *p<0.05; **p<0.01; ***p<0.001.
See Table 3 for abbreviations for regions.

Table 5. Genetic differentiation and geneflow estimates among groups.

Group 1	Group 2	*Hs*	*Ks*	*Gst*	*Fst*
FLC1					
Life forms					
Annuals	Biennials	0.842	3.889	0.024	0.037
Geographic regions					
NA	EU	0.914	4.243	0.005	0.007
NA	AP	0.840	3.796	0.019	0.015
AP	EU	0.863	4.174	0.005	0.022
FLC3					
Life forms					
Annuals	Biennials	0.947	20.421*	0.010	0.077
Geographic regions					
NA	EU	0.838	12.257	0.004	0.043
NA	AP	0.981	26.235	0.030	0.032
AP	EU	0.934**	20.223**	0.029	0.094

Significance: *, $0.01<p<0.05$, **, $0.001<p<0.01$; *Hs*=average heterozygosity among subpopulations; *Ks*= average nucleotide diversity within the two populations; *Gst*= coefficient of gene differentiation; *Fst*=equilibrium fixation index; See Table 3 for abbreviations for regions.

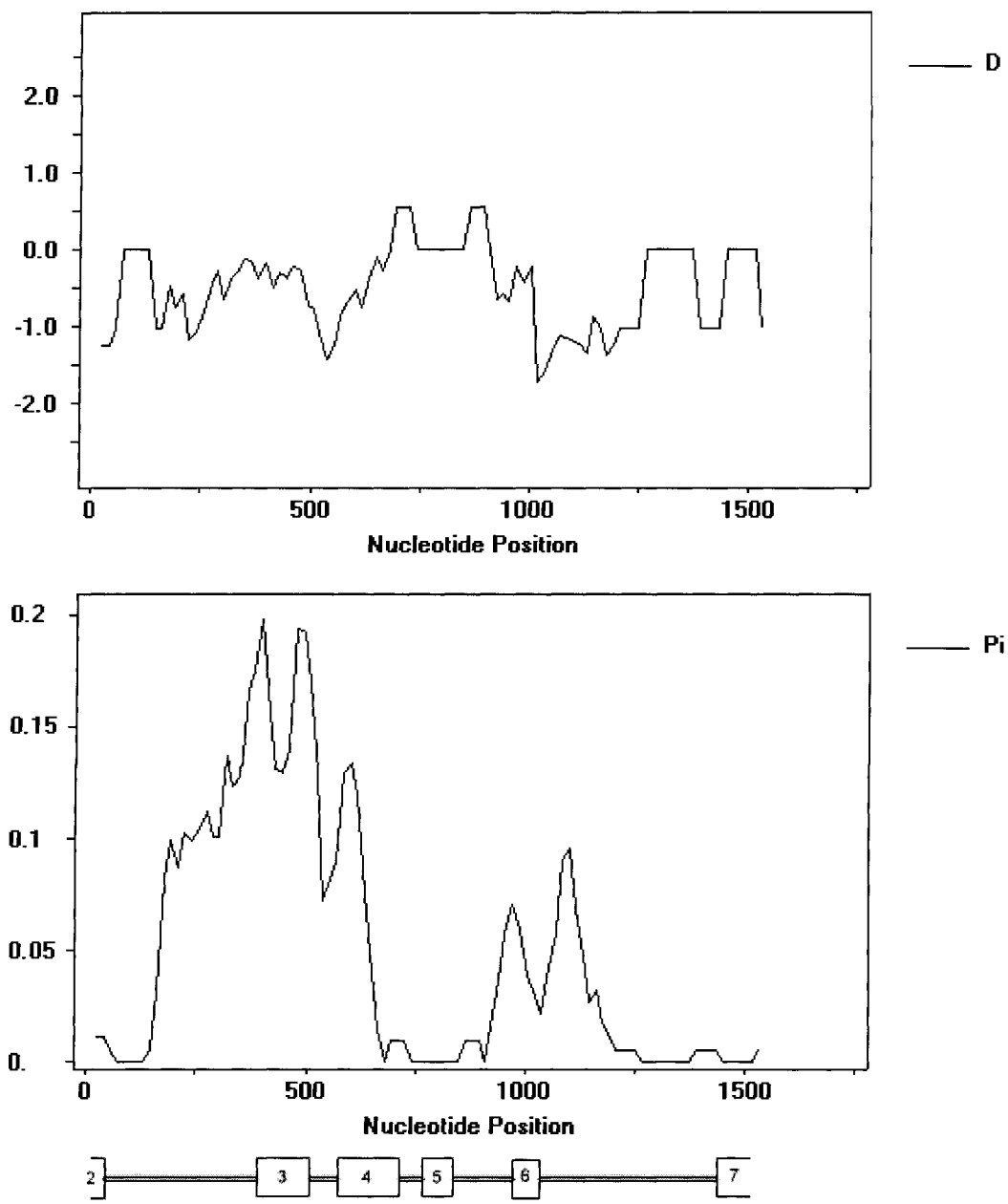


Figure 1. Plots of Tajima's *D* and nucleotide polymorphism (*Pi*) across the 1.5kb *FLC3* region surveyed in sliding window analyses. Positions of exon regions are indicated by boxes below the graphs.

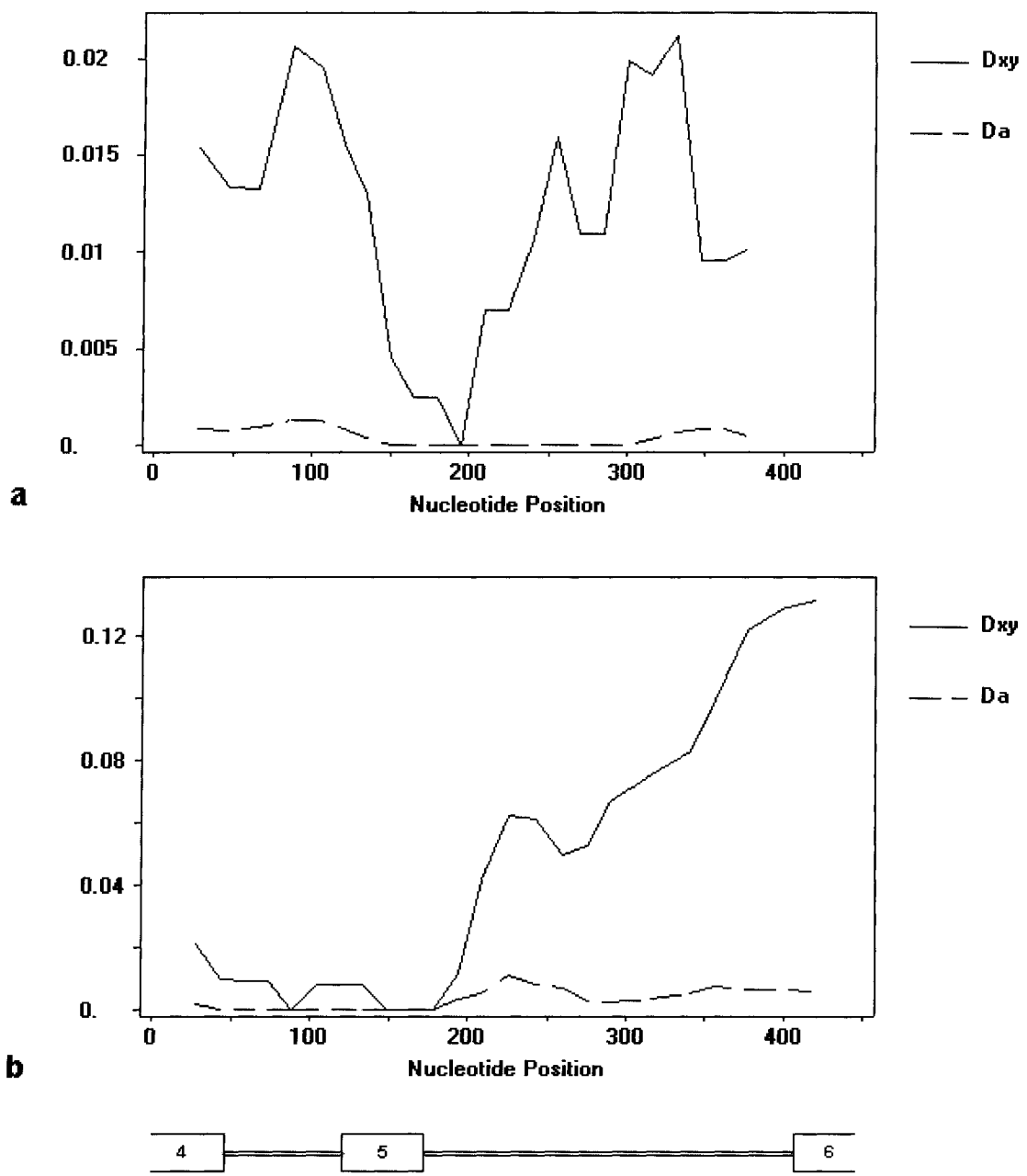


Figure 2. Plots of the interpopulation nucleotide-sequence differences (*Dxy*) and interpopulation net nucleotide-sequence divergence (*Da*) between annual and biennial groups of *B. napus* using *FLC1* (a) and *FLC3* (b) sequences. Gene region analyzed shown below the graphs, with exons indicated by boxes.

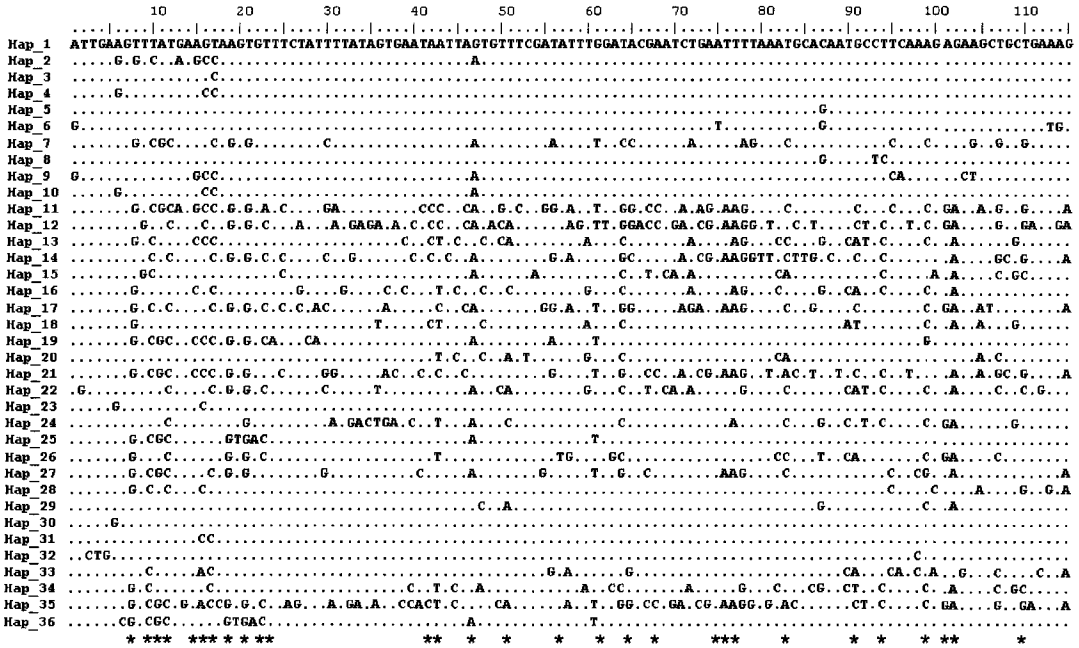


Figure 3. Polymorphisms of the 36 *FLC3* haplotypes of accessions of *B. napus*. Dots indicate the same nucleotide as observed in Hap 1. Asterisks denote positions with significant association to life form on case-control tests.

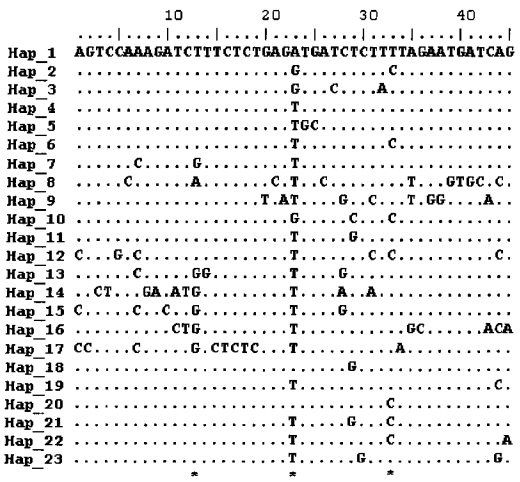


Figure 4. Polymorphisms of the 23 *FLC1* haplotypes of accessions of *B. napus*. Dots indicate the same nucleotide as observed in Hap_1. Asterisks denote positions with significant association to life form on case-control tests.

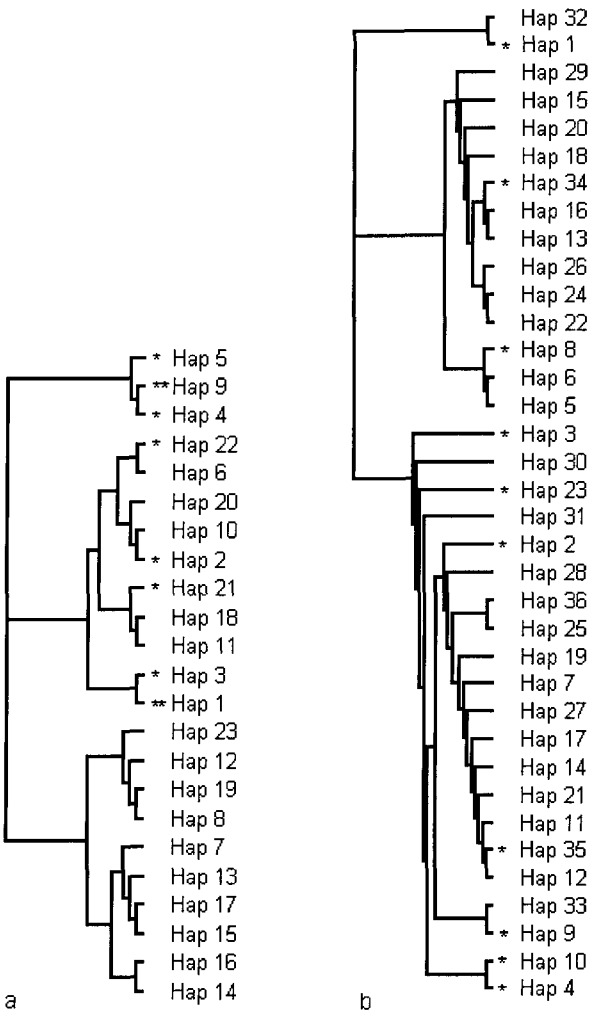


Figure 5. Rectangular cladograms showing relationships among haplotypes of *FLC1* (a) and *FLC3* (b). Each asterisk indicates a biennial accessions with observed haplotype.

CHAPTER 4. ANALYSIS OF BULKED AND REDUNDANT ACCESSIONS OF *BRASSICA* GERMPLASM USING ASSIGNMENT TESTS OF MICROSATELLITE MARKERS

A paper accepted for publication in Euphytica

Von Mark V. Cruz[1], John D. Nason[2], Richard Luhman[1*], Laura F. Marek[1], Randy C. Shoemaker[1,3], E. Charles Brummer[1] & Candice A.C. Gardner[1,4 **]

Abstract

This study was conducted to determine if *Brassica* germplasm bulks created and maintained by the USDA-ARS North Central Plant Introduction Station (NCRPIS) were made with genetically indistinguishable component accessions and to examine newly identified putative duplicate accessions to determine if they can be bulked. Using ten microsatellite primer pairs, we genotyped two bulks of *B. rapa* L. ssp. *dichotoma* (Roxb.) Hanelt comprising four accessions and three bulks of

---

[1]Department of Agronomy, Iowa State University, Ames, IA 50011

[2]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011

[3]USDA-ARS Corn Insect and Crop Genetics Research Unit, Ames, IA 50011

[4]USDA-ARS Plant Introduction Research Unit and North Central Regional Plant Introduction Station, Ames, IA 50011

** present address: Monsanto Co., Saint Louis, MO 63167*

*** author for correspondence: email: gardnerc@iastate.edu*

B. rapa ssp. *trilocularis* (Roxb.) Hanelt comprising fourteen accessions, as well as four pairs of putatively duplicate accessions of *B. napus* L. Assignment tests on ten individual plants per accession were conducted using a model-based clustering method to arrive at probabilities of accession assignment. The assignment tests indicated that one of the two bulks of *B. rapa* ssp. *dichotoma* involves genetically heterogeneous accessions. It was observed in the *B. rapa* ssp. *trilocularis* bulks that the component accessions could be differentiated into groups, with misassignments observed most frequently within groups. In *B. napus*, only one of the four pairs of putative duplicates showed significant genetic differentiation. The other three pairs of putative duplicates lack differences and support the creation of bulks. The results of the assignment tests were in agreement with cluster analyses and tests of population differentiation. Implications of these results in terms of germplasm management include the maintenance and/or re-creation of some *Brassica* germplasm bulks by excluding those accessions identified as being unique in this study.

Key words: *Brassica,* duplicate, individual assignment, genetic resources, SSRs, rapeseed, rationalization

Introduction

Brassica napus (rape) and *B. rapa* (turnip rape) are important crops that have been traditionally grown for animal fodder and as sources of inedible and edible oils (McNaughton, 1995). At present, they are also used to produce biodiesel, as well as rubber and plastics (Kimber & McGregor, 1995; Houmiel et al., 1999). *Brassica rapa* (2n=20, AA) is believed to be one of the diploid progenitor species of the amphidiploid *B. napus* (2n=38, AACC). Germplasm collections of oilseed *Brassica* in the U.S. are maintained at the USDA-ARS North Central Plant Introduction Station (NCRPIS), Ames, IA. There are currently more than 2000 accessions of oilseed *Brassica* at the NCRPIS, with the bulk comprised of *B. napus* and *B. rapa*. As part of effective management of these collections, probable duplicate accessions are routinely identified by the curator, both through the examination of passport data and via field characterization. Rationalization of germplasm collections by reducing the number of duplicate accessions is important in order to alleviate constraints on storage space and financial resources, and to simplify germplasm regeneration procedures in general (Hintum & Visser, 1995; Treuren & Hintum, 2003). For *Brassica,* regeneration costs are relatively high because many forms are biennial, requiring vernalization, and because many accessions also require pollination control in order to preserve the original genetic profile due to high levels of outcrossing. *Brassica rapa*, in particular, is an obligate outcrosser with genetic self-incompatibility (Eastham & Sweet, 2002). Similarly, *B. napus*, though often treated as autogamous, can outcross at a rate of 20-30% (Khachatourians et al., 2001; Rakow & Woods, 1987).

Bulking of *Brassica* germplasm was done by NCRPIS in 1999 and 2001. Numerous accessions of *B. rapa* in the collection received from a single source were bulked because of morphological and/or phenological similarity and a lack of definitive passport data. The morphological characterization and final bulking were performed with the help of a crop breeder. The

Brassica germplasm bulks were formed by mixing equal portions of pure live seed from each original accession. As a result, eleven new accessions of *B. rapa* ssp. *dichotoma* and six new accessions of *B. rapa* ssp. *trilocularis* were created. Each bulk has two to sixteen component accessions. The original component accessions of these *B. rapa* bulks are available; their use allowed us to test the hypothesis of accession homogeneity and determine whether the original bulks were created with genetically similar accessions.

In a similar case, apparent duplicate pairs of *B. napus* accessions were evaluated to test for similarity using molecular characterizations. These recently identified pairs have been considered for bulking because of morphological similarity, identical varietal names and origins, and consecutive genebank numbers. Similar approaches to rationalize germplasm collections by identifying and reducing the number of probable duplicates in genebanks have been conducted in other species such as *B. oleracea* L. (Hintum et al., 1996) and *Linum usitatissimum* L. (Treuren et al., 2001).

Ten microsatellite or simple sequence repeat (SSR) marker loci were used in the molecular characterization of the *Brassica* accessions. Many classes of molecular markers in *Brassica* are available, but SSRs have gained popularity because of cost effectiveness, speed, reproducibility, and especially polymorphism (Lund et al., 2003; Phippen et al., 1997; Snowdon & Friedt, 2004). Analysis of SSR data was conducted by using an assignment test to answer the question of accession homogeneity. This test uses the multilocus genotypes of representative individuals from each accession and determines if fixed differences between accessions exist. The method was first implemented by Petkau et al. (1995) and has been used successfully in population and conservation biology studies to assign individuals to specific source populations with as few as seven polymorphic marker loci (Primmer et al., 2000). The details of this method have been extensively reviewed (Cornuet & Luikart, 1996; Davies et al., 1999; Pritchard et al., 2000; Waser & Strobeck, 1998). Here, we introduce assignment tests as a tool in plant genetic resource management to assist curatorial decisions by enabling the coarse classification of, and discrimination among, putatively duplicate

accessions in *B. napus* and validate whether old bulks in *B. rapa* are composed of genetically similar accessions.

Methodology

Plant materials. Four accessions of *B. rapa* ssp. *dichotoma*, fourteen accessions of *B. rapa* ssp. *trilocularis*, and eight accessions of *B. napus* (Table 1) were selected for analysis. The *B. rapa* accessions comprise the original component accessions of old bulks whereas the *B. napus* accessions represent four pairs of putative duplicates selected from several pairs of accessions with consecutive accession numbers, identical varietal names, and similar origins as indicated in the passport data. Ten to fifteen plants per accession were grown in the greenhouse using Rootrainer™ trays (Spencer-Lemaire Ind. Ltd., Edmonton, AB) with Fafard® Canadian Growing Mix no. 2 (Conrad Fafard, Inc., Agawam, MA).

Tissue collection and DNA extraction. Leaf tissue was harvested when the plants were at the 3-4 leaf stage. Ten plants per accession were sampled. DNA was obtained from each plant using FTA® cards (Whatman Inc., Sanford, ME). Leaf extracts were collected by moderately pounding the leaves on the sample area of the cards. The cards were then allowed to dry and were stored at 24°C in a dessicator cabinet containing Drierite (W.A. Hammond Drierite Co. Ltd., Xenia, OH). To elute the DNA from the FTA® cards, 3 mm diameter discs were punched out of the sample area. Each disc was put in a 0.5 ml micro-centrifuge tube and washed with 200 µl TE (10 mM Tris, pH 8.0, 0.1 mM EDTA). After washing, the paper discs were incubated for 10-15 minutes at room temperature with 20 µl TE, and the eluted DNA used in PCR.

PCR amplification of microsatellites. PCR was performed in 96-well Microseal™ polypropylene microplates (Bio-Rad Lab. Inc., Hercules, CA). Ten SSR primer pairs from BrassicaDB (http://brassica.bbsrc.ac.uk/BrassicaDB) were used in this study (Table 2). Selection of the primers was based on presence of polymorphism observed in a previous study of *Brassica* germplasm (Cruz, 2006). The reaction mix was composed of 0.1 μl of eluted DNA, 9 μl of sterile ddH2O, 1 μl 10X PCR buffer (200 mM Tris-HCl, pH 8.4, 500 mM KCl), 0.5 μl dNTPs (2 mM), 0.3 μl MgCl (50 mM), 0.1 μl of primers (50 mM), and 0.05 μl *Taq* polymerase (5 U/μl) (Invitrogen Corp., Carslbad, CA). Thermal cycling was done using DNA Engine® (PTC-200™) thermal cyclers (Bio-Rad Lab. Inc., Hercules, CA) with the following conditions: 94°C for 2 min, then followed by 35 cycles of amplification at 94°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec, and final extension at 72°C for 4 min. The PCR products were separated in a 4.0 % agarose gel prepared in 1X TAE (40 mM Tris-acetate, 1 mM EDTA, pH 8.3) with incorporated ethidium bromide (0.46 μg/ml). Visualization of amplified products was done with a UV light box and photographed on a digital gel-documentation system.

Data analysis. Assignment tests involving putative duplicate and previously bulked accessions were conducted with Structure v.2.0 (Pritchard et al., 2000), a software package that uses a model-based clustering method and Bayesian approach to determine probabilities of assignment for samples from individual accessions. The clustering method employed by this software does not assume a particular mutation process. In conducting the simulations, each plant was tagged with its original accession number to serve as prior population information in making inferences on re-assignment. After conducting trial simulations, final program settings were used with a 30000 iteration burn-in period and 10^5 Monte Carlo Markov Chain iterations for data gathering. The null hypothesis is that the putative duplicate (or previously bulked) accessions are genetically distinct, in which case Structure should correctly group plants from the same accession. Alternatively, if putative duplicate (or

previously bulked) accessions are homogeneous, Structure will not be successful in grouping individual plants from the same accession. For diploid *B. rapa*, we also computed the probability of population differentiation between pairs of bulked accessions using GENEPOP v.3.4 (Raymond & Rousset, 1995). An experiment-wide error rate of $p<0.05$ was maintained and the significance of population differentiation values determined using critical values obtained by sequential Bonferonni adjustment (Rice, 1989). Overall relationships among accessions were visualized by using cluster analysis of genetic distances. Nei's (1972) genetic distance, which assumes a model of mutation and drift for populations of constant effective size, was computed for pairs of accessions using GenAlEx v.6 (Peakall & Smouse, 2006). Clustering of Nei's distances was performed by using the UPGMA-algorithm in NTSYS-pc v.2.2 (Rohlf, 2005).

Results

All ten microsatellite primer pairs amplified bands in the *Brassica* accessions. However, species-specific amplification was observed with the Ra2-E07, Na12-C08, and Na14-C12 primer pairs: no amplification product was obtained with Ra2-E07 in *B. napus*, with Na12-C08 in *B. rapa* ssp. *dichotoma,* or with Na14-C12 in *B. rapa* ssp. *trilocularis* (Table 2). The polymorphism information content (PIC) value of the ten microsatellite loci ranged from 0.340 (Na14-C12) to 0.793 (Ra2-F11), with an average of 0.538. An average of three bands was amplified per primer pair over all accessions in the species analyzed. Posterior probabilities derived from assignment tests on the ten plants of each accession are shown in Figures 1 to 5 together with dendrograms indicating relationships among the component accessions of bulks in *B. rapa* and the putative duplicates in *B. napus*. Specific information indicating correct assignments as well as misclassifications of plants in each accession is discussed below.

Component accessions of bulked B. rapa ssp. dichotoma

Results obtained from samples of the two pairs of original accessions that comprised the two bulks of *B. rapa* ssp. *dichotoma* are shown in Figure 1. The UPGMA dendrogram distinguished between the two bulks and grouped their component accessions according to their bulking designation. In the PI 633167 bulk, component accessions had nearly equal probabilities of assignment and the genetic distance between them was small. In contrast, in the Ames 26162 bulk, a marked difference was observed between its member accessions, with accession-specific probabilities of assignment generally greater than 0.8. Despite this difference between the two bulks, no plant was misclassified for either pair of the component accessions and tests of population differentiation indicated that component accessions in both bulks are genetically distinct ($p<0.05$).

Component accessions of bulked B. rapa ssp. trilocularis

The Ames 26168 bulk could be differentiated into groups of four (Ames 9699 [ID19], Ames 9702 [ID21], Ames 9718 [ID22], and Ames 9719 [ID23]) and two (Ames 9698 [ID18] and Ames 9701 [ID20]) component accessions (Figure 2). Also, a few of the plants in the latter two accessions had low probabilities of assignment to their source, suggesting the presence of admixture among the component accessions. Pairwise tests of population differentiation were significant for comparisons between the following component accessions: both ID18 and ID20 versus ID19, ID21, ID22, and ID23, and ID23 versus both ID19 and ID22 ($p<0.01$ in all comparisons). Misassignments were most frequent for ID21, ID22, and ID23. with almost half of the plants in each of these accessions misassigned.

The Ames 26170 bulk had probabilities of assignment that were almost equally distributed among the three component accessions with random misassignments observed in all three members of the bulk (Figure 3). Despite this similarity, the UPGMA tree of Nei's distances grouped accessions ID24 and ID26 relative to ID25, and pairwise tests of population differentiation were significant for ID25 versus ID24 and ID26 ($p<0.016$ in both cases).

In the bulk accession Ames 26171, two distinct groups were found: the first comprising Ames 9893 [ID27] and Ames 9898 [ID28], and the second Ames 9919 [ID29] Ames 9920 [ID30], and Ames 9925 [ID31] (Figure 4). Two plants were misassigned between ID27 and ID28, and three plants between ID29 and ID30. No misassignment was observed in ID31. The component accessions ID29 and ID30 were closely grouped in the UPGMA tree, had similar probabilities of assignment and did not exhibit significant population differentiation ($p=0.065$). The remaining three component accessions (Ames 9893 [ID27], Ames 9898 [ID28], and Ames 9925 [ID31]), in contrast, exhibited greater genetic distances, heterogeneous probabilities of assignment, and significant pairwise differentiation from each other and from ID29 and ID30 ($p<0.01$ in all cases). Whereas misassignments were observed in ID27 ID28, ID29 and ID30, the high probabilities of individual samples being re-assigned back into ID31 and the absence of misclassified plants suggest that this accession in particular was different from the rest of the bulk components.

Probable duplicates in B. napus

The UPGMA tree of Nei's distances grouped accessions consistent with their variety names (Figure 5). Among the four pairs of *B. napus* accessions, only the probable duplicates labeled as Aomori (PI 469724 [ID5] and PI 469725 [ID6]) showed evidence of genetic differentiation. Analysis conducted using Structure indicates that plants from these two accessions could be reassigned back to

their original source accessions with probabilities of 0.6 to 0.8. Moreover, Structure did not misclassify a single plant between the two Aomori accessions.

In the case of the other putative duplicates - Fonto (PI 469789 [ID7] and PI 469790 [ID8]), Gokstad (PI 469808 [ID9] and PI 469809 [ID10]), and Titus (PI 470046 [ID11] and PI 470047 [ID12]) - there was an average of 50% probability of assignment of samples to either accession in a pair (Figure 5). Also, several individual plants were misclassified in these pairs: two in Fonto, four in Gokstad, and four in Titus. The nearly equal probabilities of assignment, frequent misassignment, and the small genetic distance between these three accession pairs suggest that they are more likely candidates for bulking than are the two Aomori accessions.

Discussion

Microsatellites have been demonstrated to be a good source of additional information for curators to clarify relationships among phenotypically similar accessions (Phippen et al., 1997; Dean et al., 1999). In our study, differences among accessions comprising bulks and among the putative duplicates were detected using ten polymorphic microsatellite primer pairs on ten plants per accession. Our marker analyses revealed that some of the component accessions in the bulks are genetically dissimilar despite morphological and other similarities that led to their bulking. In general, as the number of component accessions in a *B. rapa* bulk increased, so did the number of genetically unique components, as observed in Ames 26168 and Ames 26171. It is recommended that these bulks be re-formulated using only genetically homogeneous accessions and excluding the unique ones.

In *B. rapa* ssp. *trilocularis*, probabilities of assignment for a few plants of Ames 9698 [ID18] and Ames 9701 [ID20] suggest that these accessions are heterogeneous, or that a low frequency of admixture might have occurred between these accessions and ID19. A similar case was observed in

Ames 9919 [ID29] and Ames 9920 [ID30], with single plants having been assigned to ID30 and ID28, respectively. Original seed lots were used in this study; admixture could have occurred before the germplasm was acquired or during seed processing and incorporation into the collection. If this is the case, identification of probable off-types in populations of these accessions prior to flower during regeneration will help prevent genetic contamination due to gene flow events.

In general, in all *B. rapa* bulks there was concordance between relationships revealed in the dendrograms, the probabilities of assignment derived from Structure, and the pairwise tests of population differentiation. The dendrograms clustered accessions with similar probability profiles, as evident in the Ames 26168, Ames 26170, and Ames 26171 bulks. The unique components of these accessions did not cluster with those that included misassigned plants. Similarly, tests of population differentiation support the relationship among accessions as indicated by the dendrogram. Accessions that are not significantly different are located in the same cluster.

Microsatellite profiles of the *B. napus* accessions differentiated putative duplicates by variety name and, with the exception of the Aomori accessions, confirmed them to be good candidates for bulking. Overall, bulking these probable duplicates is a compromise strategy that will conserve genebank resource and also increase the likelihood that unique alleles (if present) are not lost over cycles of regeneration (Hintum & Visser, 1995). As the cost of one cycle of *Brassica* germplasm regeneration in the NCRPIS is more than $500 per accession, bulking the redundant pairs also ensures cost efficiency, reducing the recurrent costs of germplasm maintenance.

It has been suggested that most probable duplicates in genebanks are 'partial' or 'common' duplicates, with only part of the alleles or genotypes in common because of regeneration and management practices. This study did not test for 'common genotypes' in the probable duplicates in *B. napus*. However, if the Aomori accessions originally came from the same source population, factors contributing to their genetic differences may include natural selection during regeneration,

contamination of seed lots, or resulting from intentional splitting of accessions into morphologically distinct parts (Hintum & Knüpffer, 1995).

Most of the approaches used in identifying duplicates rely on evaluating uniqueness of genetic profiles instead of similarity among the candidate accessions (Dean et al., 1999; Lund et al., 2003). Uniqueness has been favored over similarity because the high resolving power of molecular markers will always allow detection of some level of variation between 'identical' accessions (Treuren & Hintum, 2003). Consequently, only resource limitations restrict the discovery of variation between probable duplicates. In our study, we find that information, in the form of misassignment probabilities, genetic distance, and population differentiation, when evaluated concurrently, provides corroborative evidence that can help the crop curator decide which accessions are unique or are more likely candidates for bulking when such situations arise.

Conclusion

This study demonstrates that analyses of a modest number of microsatellites (10 loci) and molecular characterization data using assignment tests, together with more traditional methods, can be useful in verifying the homogeneity of component accessions of bulks in *B. rapa* and addressing the problem of duplicate accessions in *B. napus*. Unique accessions among *B. rapa* bulks and *B. napus* putative duplicates were identified. Use of genetic probability profiles of component accessions corroborated clustering analyses while providing additional information. Tests for population differentiation also added assurance on the uniqueness of component accessions in *B. rapa*.

Acknowledgements

This is a journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 1018, and was supported by Hatch Act and the State of Iowa. The authors wish to thank Ms. Jody Hayes and Ms. Lori Lincoln (USDA-ARS Corn Insect and Crop Genetics Research Unit) for help in the laboratory, and Dr. Mark P. Widrlechner for reviewing the draft manuscript. Mention of commercial brand names in this paper does not constitute an endorsement of any product by the U.S. Department of Agriculture or cooperating agencies.

References

Cornuet, J. M. & G. Luikart. 1996. Description and evaluation of two tests for detecting recent bottlenecks. Genetics 144:2001-2014.

Cruz, V.M.V. 2006. Molecular genetic analysis of oilseed *Brassica*: Determination of life forms and germplasm management strategies by using microsatellite markers and *FLOWERING LOCUS C* gene sequences. Ph.D. dissertation. Iowa State University, Ames, Iowa.

Davies, N., F.X. Villablanco & G.K. Roderick. 1999. Determining the source of individuals: multilocus genotyping in nonequilibrium population genetics. TREE 14(1):17-21.

Dean, R.E. J.A. Dahlberg, M.S. Hopkins, S.E. Mitchell & S. Kresovich. 1999. Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci. 39:1215-1221.

Eastham, K. & J. Sweet. 2002. Genetically modified organisms (GMOs): The significance of gene flow through pollen transfer. EEA Environmental Issue Report No. 28. European Environment Agency, Copenhagen, Denmark.

Hintum, T.J.L. & H. Knüpffer. 1995. Duplication within and between germplasm collections. I. Identifying duplication on the bases of passport data. Genet. Res. Crop Evol. 42:127-133.

Hintum, T.J.L. & D.L. Visser. 1995. Duplication within and between germplasm collections. II. Duplication in four European barley collections. Genet. Res. Crop Evol. 42:135-145.

Hintum, T.J.L., I.W. Boukema & D.L. Visser. 1996. Reduction of duplication in a *Brassica oleracea* germplasm collection. Genet. Res. Crop Evol. 43:343-349.

Houmiel, K.L., S. Slater, D. Broyles, L. Casagrande, S. Colburn, K. Gonzalez, T.A. Mitsky, S.E. Reiser, D. Shah & N.B. Taylor. 1999. Poly(beta-hydroxybutyrate) production in oilseed leukoplasts of *Brassica napus*. Planta 209: 547-550.

Khachatourians, G.G., A.K. Summer & P.W.B. Phillips. 2001. An introduction to the history of canola and the scientific basis of innovation. In: The Biotechnology Revolution in Global Agriculture. P.W.B. Phillips and G.G. Khachatourians (eds.). CABI, Wallingford, UK. pp. 33-47.

Kimber, D.S. & D.I. McGregor. 1995. The species and their origin, cultivation and world production. In: *Brassica* Oilseeds: Production and Utilization. D.S. Kimber and D.I. McGregor (eds.). CABI, Oxon, UK. pp. 1-8.

Lund, B., R. Ortiz, I.M. Skovgaard, R. Waugh & S.B. Anderson. 2003. Analysis of potential duplicates in barley gene bank collections using re-sampling of microsatellite data. Theor. Appl. Genet. 106:1129-1138.

McNaughton, I.H. 1995. Swedes and rapes – *Brassica napus* (Cruciferae). In: Evolution of Crop Plants. 2nd. Ed. J. Smartt & N.W. Simmonds (eds.). Longman Scientific & Technical, London, U.K. pp. 68-75.

Nei, M. 1972. Genetic distance between populations. American Naturalist 106:283-392.

Peakall, R. & P.E. Smouse. 2006. GenAlEx 6: Genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6: 288-295.

Petkau, D., W. Calvert, I. Stirling & C. Strobeck. 1995 Microsatellite analysis of population structure in Canadian polar bears. Mol. Ecol. 4:347-354.

Phippen, W.B., S. Kresovich, F.G. Candelas & J.R. McFerson. 1997. Molecular characterization can quantify and partition genetic variation among genebank holdings: A case study with phenotypically similar accessions of *Brassica oleracea* var. *capitata* L. (cabbage) 'Golden Acre'. Theor. Appl. Genet. 94:227-234.

Primmer, C.R., M.T. Koskinen & J. Piironen. 2000. The one that did not get away: Individual assignment using microsatellite data detects a case of fishing competition fraud. Proc. Royal Soc. London 267:1699-1701.

Pritchard, J.K., M. Stephens & P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Rakow, G. & D. Woods. 1987. Outcrossing in rape and mustard in Saskatchewan prairie conditions. Can. J. Plant Sci. 67: 147-151.

Raymond, M. & F. Rousset. 1995. GENEPOP (version 1.2): Population genetics software for exact tests and ecumenicism. J. Heredity 86:248-249.

Rice, W.R. 1989. Analyzing tables of statistical tests. Evolution 43:223–225.

Rohlf, F. J. 2005, NTSYS-pc: Numerical Taxonomy and Multivariate Analysis System, ver. 2.2. Exeter Software, Setauket, NY.

Snowdon, R.J. & W. Friedt. 2004. Molecular markers in *Brassica* oilseed breeding: Current status and future possibilities. Plant Breeding 123:1-8.

Treuren, R. & T.J.L. Hintum. 2003. Marker-assisted reduction of redundancy in germplasm collections: Genetic and economic aspects. Acta Hort. 623:139-149.

Treuren, R., L.M.J. Soest & T.J.L. Hintum. 2001. Marker-assisted rationalization of genetic resources collections: A case study in flax using AFLPs. Theor. Appl. Genet. 103:144-152.

Waser, P.M. & C. Strobeck. 1998. Genetic signatures of interpopulation dispersal. TREE 13(2):43-44.

Table 1. List of selected component accessions of five bulks in *B. rapa* ssp. *dichotoma* and *B. rapa* ssp. *trilocularis* and putative duplicates in *B. napus*.

*GH ID**	*Accession*	*Variety*	*Source***
	B. rapa ssp. dichotoma		
	Ames 26162 bulk		
14	Ames 9288	-	United States
15	Ames 9291	-	United States
	PI 633167 bulk		
16	Ames 9410	-	United States
17	Ames 9414	-	United States
	B. rapa ssp. trilocularis		
	Ames 26168 bulk		
18	Ames 9698	-	United States
19	Ames 9699	-	United States
20	Ames 9701	-	United States
21	Ames 9702	-	United States
22	Ames 9718	-	United States
23	Ames 9719	-	United States
	Ames 26170 bulk		
24	Ames 9889	-	United States
25	Ames 9895	-	United States
26	Ames 9900	-	United States
	Ames 26171 bulk		
27	Ames 9893	-	United States
28	Ames 9898	-	United States
29	Ames 9919	-	United States
30	Ames 9920	-	United States
31	Ames 9925	-	United States
	B. napus		
5	PI 469724	Aomori	South Korea
6	PI 469725	Aomori	South Korea
7	PI 469789	Fonto	South Korea
8	PI 469790	Fonto	South Korea
9	PI 469808	Gokstad	South Korea
10	PI 469809	Gokstad	South Korea
11	PI 470046	Titus	South Korea
12	PI 470047	Titus	South Korea

* Greenhouse flat number

** Source in USDA-ARS GRIN database; original *B. rapa* components from India.

Table 2. List of the ten microsatellites analyzed, band sizes observed and polymorphism information content (PIC).

SSR Name	Repeat	Primer sequence (forward, reverse)	Size range observed (~kb)	PIC
Na12-A02	$(CT)_{16}$	AGCCTTGTTGCTTTTCAACG, AGTGAATCGATGATCTCGCC	150-200	0.452
Na12-A08	$(GA)_{28}$	AACACTTGCAACTTCATTTTCC, CATTGGTTGGTGAATTGACAG	150-320	0.622
Na12-C08	$(CT)_{50}$	GCAAACGATTTGTTTACCCG, CGTGTAGGGTGATCTAGATGGG	275-350	0.537
Na14-C12	$(AG)_{17}$	CACATTTTGGTTCAATTCGG, TACGACGCTGGTTTCGATTC	190-200	0.340
Na14-D07	$(CCG)_3$	GCATAACGTCAGCGTCAAAC, CTGCGGGACACATAACTTTG	150-175	0.419
Ni4-D09	$(CT)_{25}$	AAAGGACAAAGAGGAAGGGC, TTGAAATCAAATGAGAGTGACG	170-200	0.555
Ol11-H02	$(AAC)_{18}$	TCTTCAGGGTTTCCAACGAC, AGGCTCCTTCATTTGATCCC	180-210	0.473
Ra2-E03	$(CT)_{18}$	AGGTAGGCCCATCTCTCTCC, CCAAAACTTGCTCAAAACCC	225-315	0.509
Ra2-E07	$(GA)_{19}$	ATTGCTGAGATTGGCTCAGG, CCTACACTTGCGATCTTCACC	100-170	0.677
Ra2-F11	$(CT)_{34}$	TGAAACTAGGGTTTCCAGCC, CTTCACCATGGTTTTGTCCC	190-300	0.793

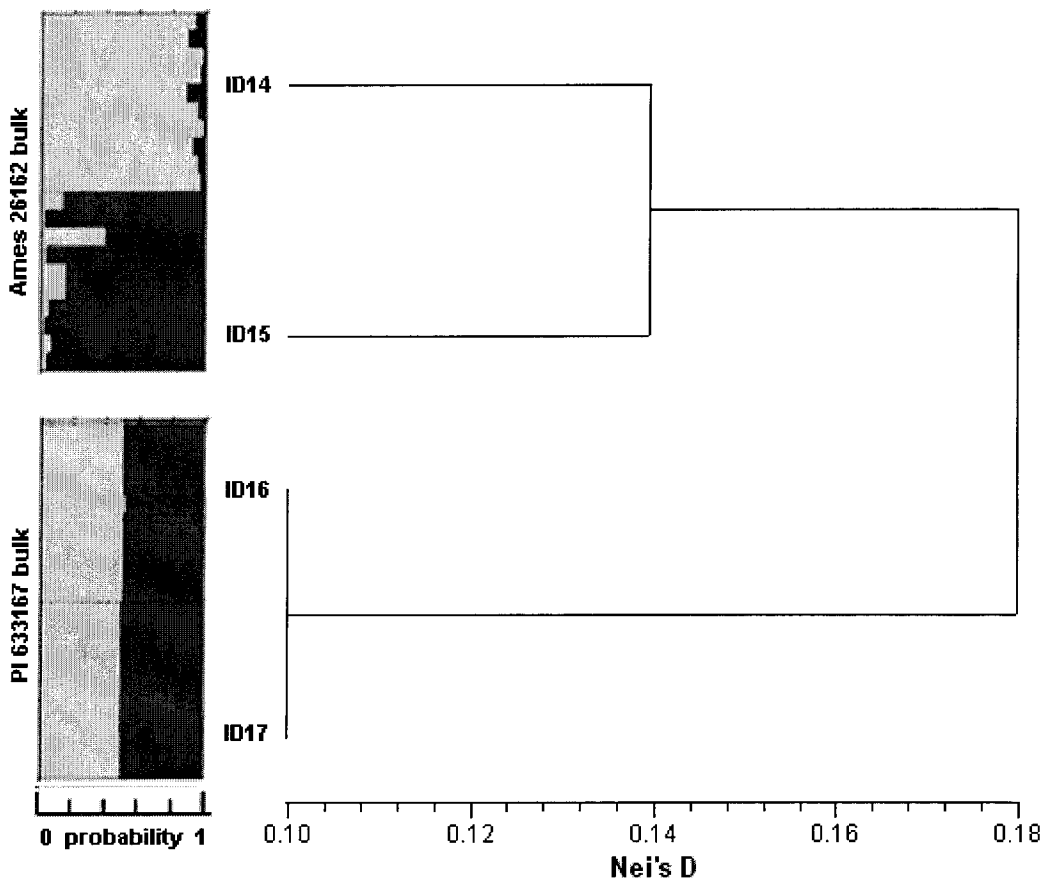


Figure 1. Relationships among the component accessions of two bulks of *B. rapa* ssp. *dichotoma* (Ames 26162 and PI 633167), and the probabilities of assignment to accessions within bulks. Pairwise comparison of differentiation between component accessions was significant between ID14 and ID15 (p=0.002) and between ID16 and ID17 (p=0.004).

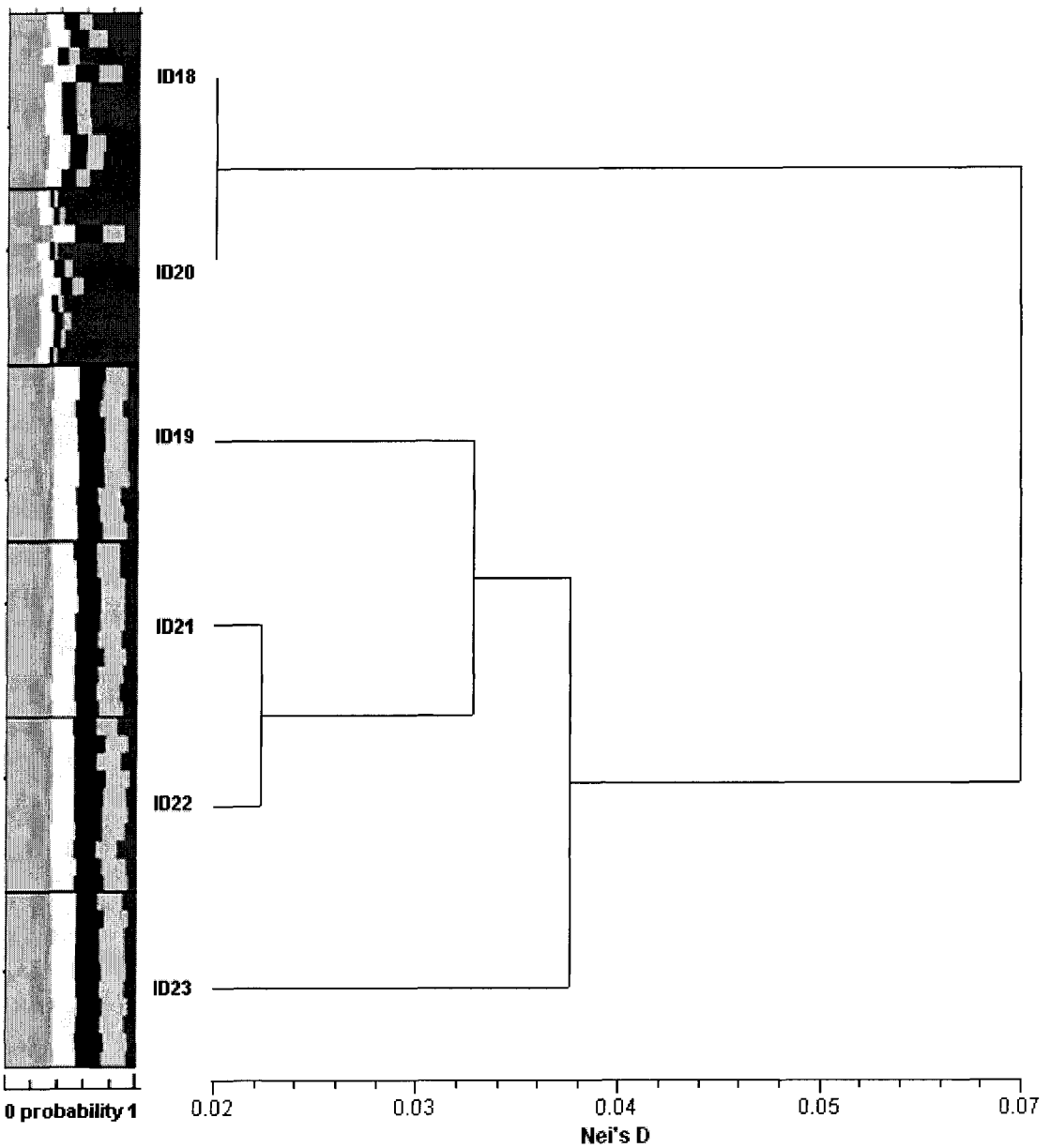


Figure 2. Relationships among the component accessions of Ames 26168 bulk (*B. rapa* ssp. *trilocularis*) and the probabilities of assignment of individual samples to each accession. Colors portrayed (L to R) correspond to probabilities of assignment to accessions ID23, ID22, ID21, ID20, ID19, and ID18, respectively. Pairwise comparison of differentiation between component accessions were non-significant in ID18 and ID20 (p=0.173), ID19 and ID22 (p=0.058), ID19 and ID21 (p =0.737), ID21 and ID22 (p=0.672), and ID21 and ID23 (p=0.185); all other comparisons have significant p-values (p<0.01).

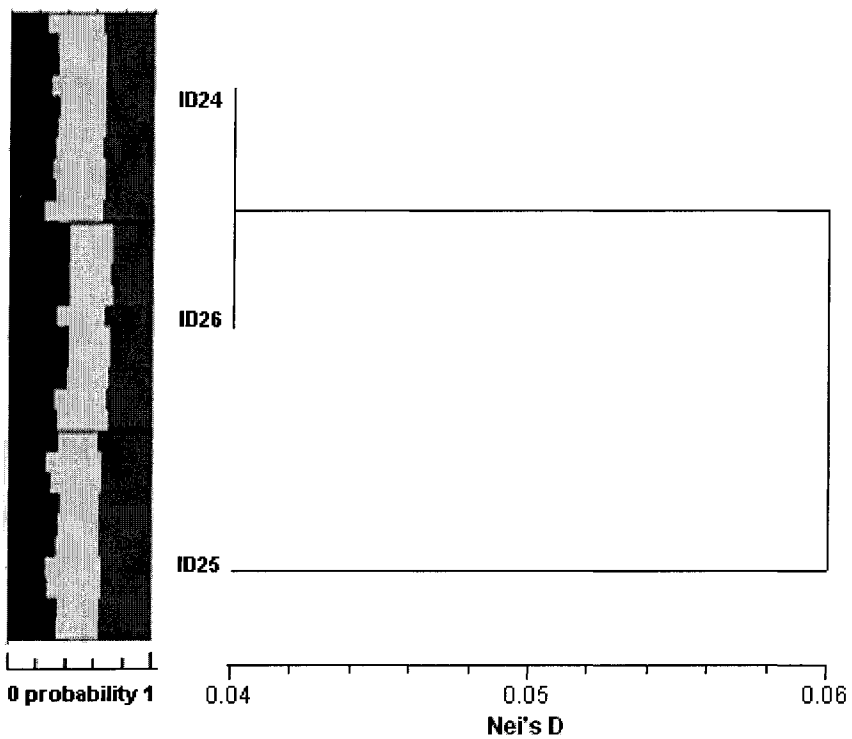


Figure 3. Relationships among the component accessions of Ames 26170 bulk (*B. rapa* ssp. *trilocularis*) and the probabilities of assignment of the samples within each accession. Colors (L to R) correspond to probabilities of assignment to accessions ID26, ID25, and ID24, respectively. Pairwise comparison of differentiation between component accessions was non-significant between ID24 and ID26 (p=0.543) and highly significant for all other comparisons (p<0.016).

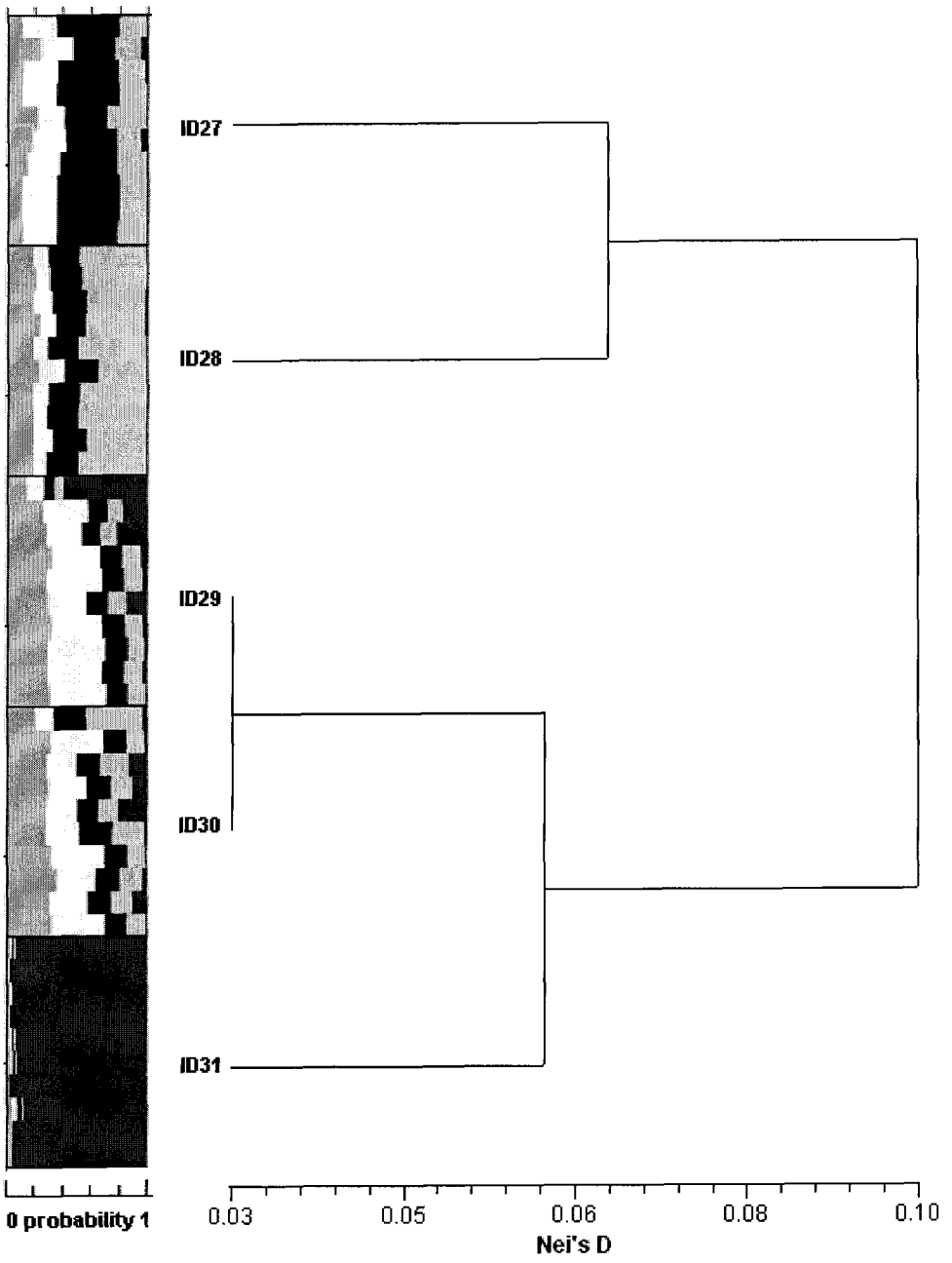


Figure 4. Relationships among the component accessions of Ames 26171 bulk (*B. rapa* ssp. *trilocularis*) and the probabilities of assignment of the samples within each accession. Colors (L to R) correspond to probabilities of assignment to accessions ID31, ID30, ID29, ID28, and ID27, respectively. Pairwise comparison of differentiation between component accessions was non-significant between ID29 and ID30 (p=0.065) and highly significant for all other comparisons (p<0.01).

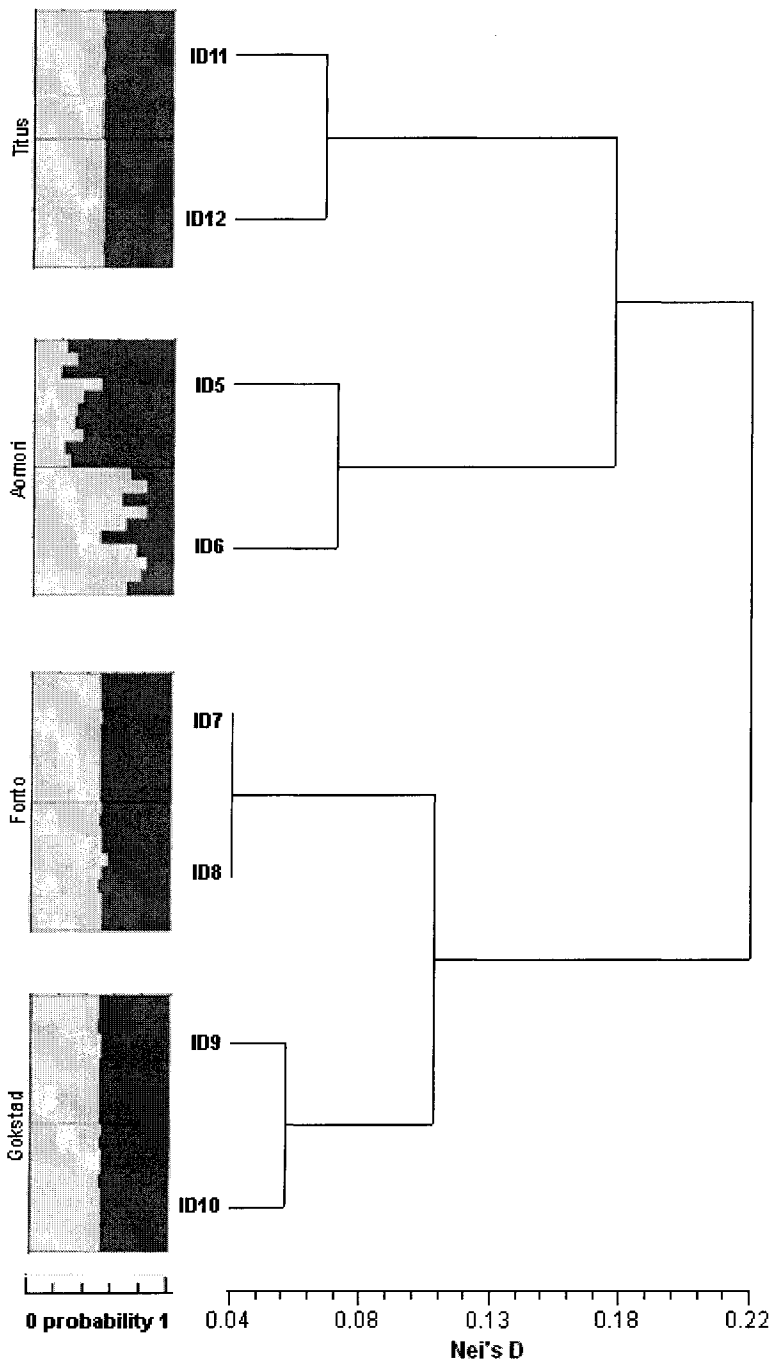


Figure 5. Dendrogram constructed from Nei's genetic distance showing the relationships among the putative *B. napus* duplicates indicated in Table 1. Probabilities of assignment within an accession are indicated to the left. Probabilities of 0.5 indicate genetic homogeneity of putative duplicates, whereas consistently skewed probabilities (see Aomori) indicate genetic heterogeneity of putative duplicates.

CHAPTER 5. MEASURING THE EFFECTIVENESS OF ISOLATION OF *BRASSICA NAPUS* L. ACCESSIONS DURING CAGED GERMPLASM REGENERATION

A paper submitted for publication in Genetic Resources and Crop Evolution

Von Mark V. Cruz[1], Charlie L. Rife[2*], John D. Nason[3],

E. Charles Brummer[1] & Candice A.C. Gardner[1,4**]

Abstract

At the North Central Regional Plant Introduction Station (NCRPIS), screen cages and pollinator insects are used during *Brassica* germplasm regeneration to control outcrossing among accessions and enable the plants inside a cage to intermate. Previous reports on *B. napus* indicated that both insect and wind pollination occur under open-field conditions. This study was conducted to determine

---

[1]Department of Agronomy, Iowa State University, Ames, IA 50011

[2]Department of Agronomy, Kansas State University, Manhattan, KS 66506

[3]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, IA 50011

[4]USDA-ARS, Plant Introduction Research Unit and North Central Regional Plant Introduction Station, Ames, IA 50011

[*] *present address: Blue Sun Biodiesel, Westminster, CO 80234*

[**] *author for correspondence: email: gardnerc@iastate.edu*

whether current germplasm-regeneration methods preclude viable pollen of external origin from entering adjacent caged plots, contaminating the genetic integrity of *B. napus* regenerations. A dominant herbicide-resistance trait, Roundup-Ready® (RR), was chosen to detect the frequency of progeny resulting from cross pollination of plants in adjacent cages by an external pollen source because of ease of bioassay methods. A caged field was established where a non-transgenic *B. napus* variety was planted in two replicates surrounded by a transgenic RR isoline. Pollen escape was detected and quantified by using pollen traps placed outside the cages during the first two weeks of flowering. There was no significant correlation between pollen count and wind speed or direction. Bioassays of the progenies from the non-transgenic variety indicated that pollen flow and contamination between cages did occur, but at a very low frequency, 0.01%. These results indicate that the regeneration methods employed at the NCRPIS for *B. napus* are suitable for controlling external contamination.

Key words: canola, genebank, herbicide resistance, pollination, pollen flow, rapeseed

Introduction

Gene flow, which is the successful movement of genes among populations by mating or migration of seeds or other propagules, has gained much interest in agriculture in recent years because of the widescale adoption of transgenic crops and the concern of transgene escape into the wild (James, 2004; Messeguer, 2003; Stewart et al., 2003). *Brassica napus* (rapeseed), together with maize and sugar beets, have been identified among the species for which cross pollination and transgene escape are concerns (Treu & Emberlin, 2000).

In the conservation of plant genetic resources, pollination-control methods to preserve the genetic integrity of regenerations include pollination by hand and the use of screen or fabric cages

with insect pollinators. Failure of these methods to control contamination by any pollen source, regardless of whether the pollen is transgenic or non-transgenic, can compromise the integrity of the collections, which limits their utility to researchers.

Screened cages are widely used to enclose growing plants during germplasm regeneration, as in the case of *Capsicum* in New Mexico (Bosland, 1993), and ornamentals, *Cucumis*, *Helianthus*, and *Brassica* at the NCRPIS (NCRPIS, 2002). The purpose of the cages is to maintain the genetic integrity of each accession by preventing cross contamination among adjacent accessions in the field while allowing enclosed plants to intermate with the help of insect pollinators. The NCRPIS cage system has been found effective in maintaining accession germplasm identity in *Helianthus annuus* and *Cucumis sativus* (Widrlechner et al., 1992; Wilson, 1989), but no similar study has been conducted on this system for *Brassica*.

Brassica napus has a mixed mating system where cross pollination occurs at an average of 30% (Warwick & Miki, 2004). The rate of outcrossing can vary among populations and locations, ranging from 5% to 47% (Becker et al., 1992: Lavigne et al., 1998; Rakow & Woods, 1987). Cross pollination was thought to be effected only by insects in *B. napus*. However, Timmons et al. (1995) and Wilkinson et al. (2003) indicated that wind plays a significant role in its pollination even over long distances. Substantial amounts of *Brassica* pollen are released into the air during flowering and can remain airborne (McCartney and Lacey, 1991). Lewis (1983) and Sarkissian and Harder (2001) have reported that the diameter of *Brassica* pollen ranges from 24 to 27 μm. Yet popular screens used for caging have a low resistance to wind (Bell and Baker, 2000) and mesh openings of 280 to more than 1000 μm (BioQuip, 2005). Considering the small size of *B. napus* pollen relative to the mesh holes of the cage screens, it is conceivable that pollen could pass through screens and effect cross pollination. If pollen can pass through the screens, the probability of cross pollination may be inversely proportional to distance between cages during regeneration. In addition, higher outcrossing rates could result for self-incompatible *Brassica* accessions (Hansen et al., 2003).

The potential for contamination is a concern in maintaining the integrity of accessions especially those of outcrossing species during germplasm regeneration in genebanks (Ramanatha Rao & Hodgkin, 2002). Because the regenerated plants often represent a subsample of the accession, the impact of gene-flow events can be magnified. With small population sizes, gene-flow events can occur at higher rates with stronger effects (Goodell et al., 1997). This study was designed to determine if *B. napus* pollen contamination occurs under the standard cage-regeneration protocols currently employed at the NCRPIS when using Lumite® screening and to determine the rate and significance of gene-flow events during regeneration.

Materials and Methods

Plant materials. Two varieties of *B. napus* with overlapping flowering times were selected: a conventional hybrid (Hyola 401), and its Roundup-Ready® (RR) isoline (Hyola 357RR), kindly provided by the Interstate Seed Company, West Fargo, ND. Glyphosate-herbicide resistance provided a dominant-trait model for which a bioassay can be easily conducted.

Upon receipt, 200 seeds from each variety were germinated in the greenhouse and tested for RR-trait purity following a screening procedure modified from Reboud (2003). Glyphosate herbicide was applied at 1080 g/ha at 300 L/ha during the two leaf stage and again two weeks after the initial application.

Field layout and planting. Two replications each containing nine plots were established on the Bennett Farm (41°59'36"N, 93°41'22"W) of Iowa State University (ISU) (Figure 1a). In each group, the center plot was planted with Hyola 401 and the surrounding plots with Hyola 357RR. The design maximized the opportunity for the non-RR center plots to capture pollen from an RR source given any possible wind direction. Direct seeding was done by using a V-belt push planter (Allan

Machine Company, Nevada, IA) following the row and plant distances from the NCRPIS protocol for *Brassica* planting (NCRPIS, 2002). Non-RR plots were sown first to prevent possible contamination of the non-RR plots with RR seeds, and the planter was thoroughly cleaned prior to planting each variety. The plots were thinned to 100 plants in 2 rows (50 plants per 5.5 m row) to give a density of 18 plants/m^2. To verify that no admixing of seeds occurred during planting, all the plants in the non-RR plots were checked for the presence of RR plants before flowering. Lateral-flow membrane strips (QuickStix™ Kit for Roundup Ready® Canola Leaf & Seed, Envirologix, Portland, ME) (one strip per two plants) were used to detect the expression of *cp4 5-enolpyruvyl-shikimate-3-phosphate synthase* (*EPSPS*) protein in leaf tissues and identify transgenic plants.

Screen cages, insect pollinators and plot care. The plots were covered with screen cages, each 1.5 m tall × 1.5 m wide × 6.0 m long before the plants began flowering (Figure 1b). The screen cages used at the NCRPIS were previously described by Widrlechner et al. (1996). The screen cages were spaced 4.5 meters apart in the field. The Lumite® screen material of the cages has a mesh size of 7 × 6 per cm^2. The average size of the screen mesh hole is 1394 ± 49 (s.e.) μm.

A colony of honeybees (*Apis mellifera*) in a 'nuc box' (16.8 cm × 50.8 cm × 27 cm) was placed in each of the non-RR cages when the first flower appeared and contained inside until the end of flowering. No pollinator insects were placed in the RR cages to ensure that pollen escape from the RR cages would be due to wind action and not due to escaped bees. No feral bees were observed during the flowering period.

Field fertility levels were high and no fertilizer was applied. No aphids or thrips were observed inside the cages during the season. Acephate insecticide was sprayed once during the plants juvenile stage to control alfalfa loopers (*Autographa* sp.).

Pollen and wind measurements. The direction and speed of wind during the flowering period were recorded at 10-minute intervals by a portable weather station placed in the center of the field. Simple pollen traps were constructed following the method of Hoekstra (1965). The pollen traps have two circular sticky surfaces, each with an area of 38.5 mm^2. The static traps were positioned in four locations in the field (2 m from each cage) at the start of flowering (Figure 1a).

Pollen grain sizes were determined from flowers harvested from the RR and non-RR plants. Measurements were conducted using a Z2™ Particle Count and Size Analyzer (Beckman Coulter, Inc., Fullerton, CA) with a 100 μm aperture. Size measurements were obtained from 1 ml of pollen collected in Isoton® II Diluent (Beckman Coulter, Inc., Fullerton, CA).

Detecting gene flow. Seeds were harvested from the cages of RR and non-RR varieties while the screens were still on. Seeds from the non-RR cages were sown in the greenhouse in 1020 plastic nursery flats (T.O. Plastics Inc., Minneapolis, MN) with Fafard® Canadian Growing Mix no. 2 (Conrad Fafard, Inc., Agawam, MA). The frequency of detection of RR progeny from the cages of non-RR plants provided an estimate of the rate of gene flow between accessions resulting from pollen escape and successful pollination. Out of the 9500 seeds planted, 9220 seedlings reached the two leaf stage and were screened for resistance. Glyphosate herbicide was applied at 1080 g/ha at 300 L/ha during the two leaf stage and re-spraying was done two weeks after the initial application. Seeds germinated from both the original samples of Hyola 401 and Hyola 357RR served as susceptible and resistant controls. Contamination frequency (gene flow) was estimated by calculating the proportion of resistant seedlings, 14 days post-herbicide application. The results were validated by testing the surviving plants for expression of *cp4 EPSPS* protein by using lateral-flow membrane strips.

Results and Discussion

Flowering began 45 days after planting and continued for more than three weeks. The RR plants started flowering four days earlier than the non-RR plants. The mean pollen size of RR plants was determined to be 29.20 ± 0.07 (s.e.) μm and the non-RR plants 29.92 ± 0.13 (s.e.) μm (Figure 2). These sizes are larger than those reported by Lewis (1983) or observed by Sarkissian and Harder (2001) in *B. rapa*. There was no significant differences between the pollen sizes of the non-RR and RR varieties (p<0.18).

Clumps of *B. napus* pollen (two or more grains) were detected in sticky traps outside cages after the initiation of flowering. Clumping, or aggregation, is due to pollen stickiness and has been previously reported; it is believed to be an adaptation to favor dispersal by animals (Creswell et al., 2004). The highest frequency of pollen detection was observed during the second week of monitoring when the plants were at 50% flowering (Figure 3). There was no significant correlation between the quantity of pollen trapped versus wind speed (r=0.20; $t_{0.05, 23}$=0.98, t_{crit}=1.71), wind gust (r=0.22; $t_{0.05, 23}$=1.08, t_{crit}=1.71) or wind direction (r=-0.06, $t_{0.05, 23}$=0.29, t_{crit}=1.71).

Mature pods were harvested 83 days after planting. Seed set on caged plants of the RR variety in the absence of insect pollinators was less than seed set from caged non-RR plants with bee pollinators, but not significantly different (p<0.33). The results agree with previous studies, which found out that there was comparable seed set between caged plants with pollinators and those without pollinators in *B. napus* (Free and Nuttall, 1968; Lerin, 1982). A previous study at the NCRPIS indicated that significant seed yield increases were obtained in *B. rapa,* but not in *B. napus* in controlled-pollination cages using various bee species (Wilson et al., 1999). Pollinator insect activity may be more important in *B. rapa* and other *Brassica* species with higher expression of self incompatibility (Thomas, 2003).

At present, wind is believed to be a major initiator of cross or self-pollination in *B. napus*, and bees are unlikely to have substantial effects on yield (Lerin, 1982; Thomas, 2003). Experiments conducted on *Brassica* grown in the still air of a glasshouse resulted to poor yields (Williams, 1978). In our experiment, seed set occurred due to either self or sib fertilization. Without pollinators, this could have been effected by the wind, enabling the transfer of viable pollen to receptive stigmata due to shaking of the plants (Williams, 1978). Results of wind tunnel experiments on *B. napus* flowers do not reject the possibility that stigmata can capture airborne pollen (Creswell et al., 2004).

Brassica napus is an allotetraploid (4x, 2n=38) and all currently commercialized transgenic canola varieties of this species carries two dominant genes for Roundup resistance, a copy each of the *cp4 EPSPS* gene from *Agrobacterium* sp. and a modified glyphosate oxidoreductase gene (*goxv247*) based on *gox* from *Ochrobactrum anthropi* (Warwick & Miki, 2004). Given the two copies of the resistance genes, the genotype for the tetraploid hybrid can be noted as RRxx (x denoting the null allele). The gametic output for an allotetraploid is expected to follow that of a diploid (Comai, 2005). In *Brassica*, chromosome pairing was reported to be genetically controlled (similar to the *Ph1* mechanism in wheat) and varies among plant varieties (Jenczewski et al., 2003). In the complete absence of preferential chromosome pairing, the pollen (n=2x) from the transgenic line would have three possible genotypes: RR, Rx and xx, occurring with probabilities of 0.25, 0.50, and 0.25, respectively. Due to these possible genotypes and associated probabilities, there remains a small amount of undetectable pollen flow. The proportion of pollen that has the null genotype cannot be determined empirically following our experimental design and procedure.

Two large seed samples from the non-RR harvest were planted in the greenhouse, resulting in 9,220 seedlings (rep1=4,523, rep2=4,697), which were sprayed with glyphosate to test for resistance. A single plant in replicate 1 tested positive for herbicide resistance, and the subsequent lateral-flow strip test verified the presence of the transgene. These results correspond to a contamination rate of 0.01% (1:9,200). If one accounts for bias created by undetected gene flow from the xx pollen, this

estimate would only increase by 0.003% at most, or equal 0.013%. This estimate is within the genetic purity limit established by the Association of Official Seed Certifying Agencies (AOSCA), which allows a maximum of 0.25% contamination in certified canola seedlots and 0.05% for breeder and foundation seedlots (Friesen et al., 2003).

The contamination we recorded in the cage experiment is comparable to the hybridization rate observed by Scheffler et al. (1995) in *B. napus* test plots planted 200m apart in an open field. It is also much lower than the observed contamination in small plots planted 1.5m apart without cages (Funk et al., 2006). A zero percent geneflow between accessions is desirable, but difficult to attain in a cross-fertilizing crop such as *Brassica*. Insect exclusion cages, with coarse or fine weave netting, were reported by Ramsay et al. (2003) to reduce pollination in their study but also did not completely prevent it. We are confident that given the observed level of detected outcrossing (<0.02%), the current screened-cage production system utilized at the NCRPIS sufficiently isolates each accession to preclude contamination due to inadvertent pollen flow during regeneration.

Conclusion

This study demonstrated that pollen from a *B. napus* variety grown in screened cages in the field during germplasm regeneration can pass through the screen mesh currently at use by the NCRPIS and cause contamination across cages at a low frequency. The contamination occurred below the current AOSCA limit set for maintaining genetic identity giving us confidence that the current method of regenerating *B. napus* in screen cages is functional. It is yet to be determined whether switching to a finer screen or increasing the distance between cages could completely eliminate pollen flow, and if levels of contamination vary across the more than 20 species of *Brassica* being conserved at the NCRPIS.

Acknowledgements

This is a journal paper of the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, Project No. 1018, and was supported by Hatch Act and the State of Iowa, and by the U.S. Department of Agriculture (USDA), Agricultural Research Service, Plant Introduction Research Unit. We wish to thank Mr. Jim Johnson (Interstate Seed Co) for providing the canola seeds, Mr. Mike Stahr (ISU Seed Science Center) for the test strips, Dr. Madan Bhattacharyya, Dr. Mark Westgate, Ms. Maria Hartt (ISU Agronomy Department), and Dr. Barbara Bingaman (NCRPIS, ISU Agronomy Department) for technical assistance and the use of laboratory and field equipments, and Mr. Larry Lockhart, Mr. Lloyd Crim, and Mr. Brian Buzzell (NCRPIS, ISU Agronomy Department) for help in land preparation, maintenance, and cage assembly. We are also grateful to Dr. Mark Widrlechner for reviewing the manuscript and the Fulbright Program of the U.S. Department of State's support of VMVC during this study. Mention of commercial brand names in this paper does not constitute an endorsement of any product by the U.S. Department of Agriculture or cooperating agencies.

References

Becker, H.C., R. Carle & S.S. Han. 1992. Environmental variation for outcrossing rates in rapeseed (*Brassica napus*). Theor. Appl. Genet. 84:303-306.

Bell, M.L. & J.R. Baker. 2000. Comparison of greenhouse screening materials for excluding whitefly (Homoptera: Aleyrodidae) and thrips (Thysanoptera: Thripidae). J. Econ. Entom. 93:800-804.

BioQuip. 2005. Equipment Catalog. BioQuip® Products, Inc., Rancho Dominguez, CA. 203 pp. http://www.bioquip.com.

Bosland, P.W. 1993. An effective plant field cage to increase the production of genetically pure chile (*Capsicum* spp.) seed. HortScience 28:1053.

Comai, L. 2005. The advantages and disadvantages of being polyploid. Nature Rev. Genet. 6:836-846.

Cresswell, J.E., T.W. Davies, M.A. Patrick, F. Russell, C. Pennel, M. Vicot & M. Lahoubi. 2004. Aerodynamics of wind pollination in a zoophilous flower, *Brassica napus*. Functional Ecology 18:861-866.

Free, J.B. & P.M. Nuttall. 1968. The pollination of oilseed rape (*Brassica napus*) and the behaviour of bees on the crop. J. Agric. Sci. 71:91-94.

Freisen, L.F., A.G. Nelson & R.C. Van Acker. 2003. Evidence of contamination of pedigreed canola (*Brassica napus*) seedlots in western Canada with genetically engineered herbicide resistance traits. Agron. J. 95:1342-1347.

Funk, T., G. Wenzel & G. Schwarz. 2006. Outcrossing frequencies and distribution of transgenic oilseed rape (*Brassica napus* L.) in the nearest neighborhood. Europ. J. Agronomy. 24:26-34.

Goodell, K., D.R. Elam, J.D. Nason & N.C. Ellstrand. 1997. Gene flow among small populations of a self-incompatible plant: An interaction between demography and genetics. Am. J. Botany 84:1362-1371.

Hansen, L.B., H.R. Siegismund & R.B. Jorgensen. 2003. Progressive introgression between *Brassica napus* (oilseed rape) and *B. rapa.* Heredity 91:276-283.

Hoekstra, P.E. 1965. A simple, inexpensive pollen trap. U.S. Forest Service Research Note SE-40. U.S. Department of Agriculture Forest Service. Asheville, North Carolina.

James, C. 2004. Preview: Global status of commercialized biotech/GM crops: 2004. ISAAA Brief 32. ISAAA, Ithaca, NY.

Jenczewski, E., F. Eber, A. Grimaud, S. Huet, M.O. Lucas, H. Monod & A.M. Chevre. 2003. *PrBn*, a major gene controlling homeologous pairing in oilseed rape (*Brassica napus*) haploids. Genetics 164:645-653.

Lavigne, C., E.K. Klein, J. Pierre, B. Godelle & M. Renard. 1998. A pollen-dispersal experiment with transgenic oilseed rape. Estimation of the average pollen dispersal of individual plant within a field. Theor. Appl. Genet. 96:886-896.

Lerin, J. 1982. Effets de la pollinisation entomophile sur le colza dans une experience en cage (Effects of pollination on rape in a cage experiment) (Abstract). Agronomie 2:249-255.

Lewis, W.H. 1983. Airborne and allergenic pollen of North America. The Johns Hopkins University Press, Baltimore, MD. pp. 150-151.

McCartney, H.A. & M. E. Lacey. 1991. Wind dispersal of pollen from crops of oilseed rape (*Brassica napus* L.). J. Aerosol. Sci. 22:467-477.

Messeguer, J. 2003. Gene flow assessment in transgenic plants. Plant Cell, Tissue and Organ Culture 73:201-212.

NCRPIS. 2002. North Central Regional Plant Introduction Station (NCRPIS) Operations Manual. Ames, IA. 76 pp.

Rakow, G. & D. Woods. 1987. Outcrossing in rape and mustard in Saskatchewan prairie conditions. Can. J. Plant Sci. 67: 147-151.

Ramanatha Rao, V. & T. Hodgkin. 2002. Genetic diversity and conservation and utilization of plant genetic resources. Plant Cell, Tissue and Organ Culture 68:1-19.

Ramsay, G., C. Thompson & G. Squire. 2003. Quantifying landscape-scale gene flow in oilseed rape. Final Report of DEFRA Project RG0216. Department for Environment, Food and Rural Affairs (DEFRA), London, U.K. 50 pp.

Reboud, X. 2003. Effect of a gap on gene flow between otherwise adjacent transgenic *Brassica napus* crops. Theor. Appl. Genet. 106: 1048-1058

Sarkissian, T.S. & L.D. Harder. 2001. Direct and indirect responses to selection of pollen sizes in *Brassica rapa* L. J. Evol. Biol. 14:456-468.

Scheffler, J.A., R. Parkinsson & P.J. Dale. 1995. Evaluating the effectiveness of isolation distances for field plots of oilseed rape (*Brassica napus*) using a herbicide-resistance gene as a selectable marker. Plant Breeding 114:317-321.

Stewart, C.N Jr., M.D. Halfhill & S.I. Warwick. 2003. Transgene introgression from genetically modified crops to their wild relatives. Nature Rev. Genet. 4:806-817.

Thomas. P. 2003. Canola Grower's Manual. Canola Council of Canada. Winnipeg, Manitoba. pp. 307 and 1017a.

Timmons, A.M., E.T. O'Brien, Y.M. Charters, S.J. Dubbels & M.J. Wilkinson. 1995. Assessing the risks of wind pollination from field of genetically modified *Brassica napus ssp. oleifera.* Euphytica 85:417-423.

Treu, R. & J. Emberlin. 2000. Pollen dispersal in the crops maize (*Zea mays*), oil seed rape (*Brassica napus* ssp. *oleifera*), potatoes (*Solanum tuberosum*), sugar beet (*Beta vulgaris* ssp. *vulgaris*) and wheat (*Triticum aestivum*). http://www.soilassociation.org.

Warwick, S. & B. Miki. 2004. Herbicide resistance. In: Biotechnology in Agriculture and Forestry 54: *Brassica*. E.C. Pua & C.J. Douglas (eds.). Springer-Verlag, Heidelberg, Germany. pp. 273-290.

Widrlechner, M.P., C.A. Abel & R.L. Wilson. 1996. Ornamental seed production in field cages with insect pollinators. Combined Proceedings International Plant Propagators' Society 46: 512-516.

Widrlechner, M.P., L.D. Knerr, J.E. Staub & K.R. Reitsma. 1992. Biochemical evaluation of germplasm regeneration methods for cucumber, *Cucumis sativus* L. FAO/IBPGR Plant Genetic Res. Newsletter 88/89:1-4.

Wilkinson, M.J., L.J. Elliot, J. Allainguillaume, M.W. Shaw, C. Norris, R. Welters, M. Alexander, J. Sweet & D.C. Mason. 2003. Hybridization between *Brassica napus* and *B. rapa* on a national scale in the United Kingdom. Science 302:457-459.

Williams, I. 1978. The pollination requirements of swede rape (*Brassica napus* L.) and of turnip rape (*Brassica campestris* L.). J. Agric. Sci. 91:343-348.

Wilson, R.L. 1989. Minimizing the extraneous transfer of sunflower pollen by honey bees (Hymenoptera: Apidae) in field cages. J. Kansas Ent. Soc. 62:387-391.

Wilson, R.L., C.A. Abel & R.L. Luhman. 1999. Comparing three bee species for controlled pollination of selected Brassicaceae. Jour. Iowa Acad. Sci. 106:1-3.

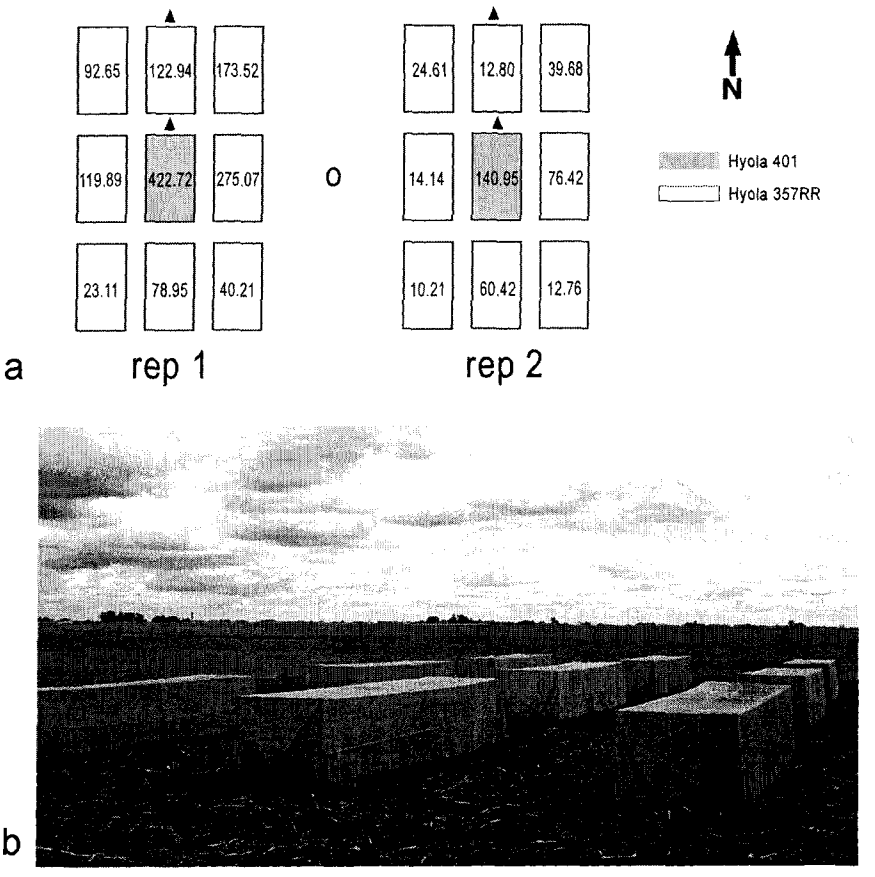


Figure 1. a) Field layout showing the position of the Roundup-Ready® (Hyola 357RR) and non-resistant (Hyola 401) *B. napus* varieties, 'o' indicates the location of the field weather station and '▲' the pollen traps. Numbers designate quantity of seeds (grams) harvested from the cage (mean 100-seed weight= 0.28g); b) view of the screen cages from one replicate.

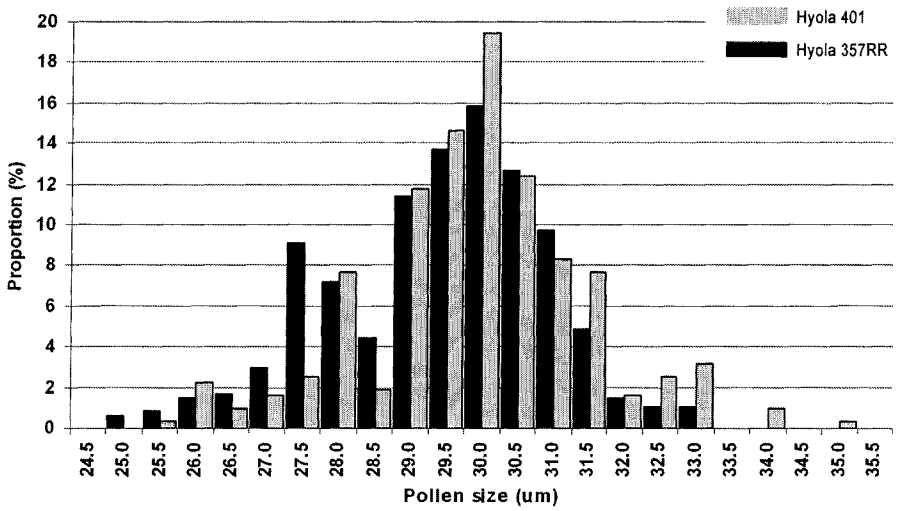


Figure 2. Pollen sizes observed in Hyola 401 and Hyola 357RR.

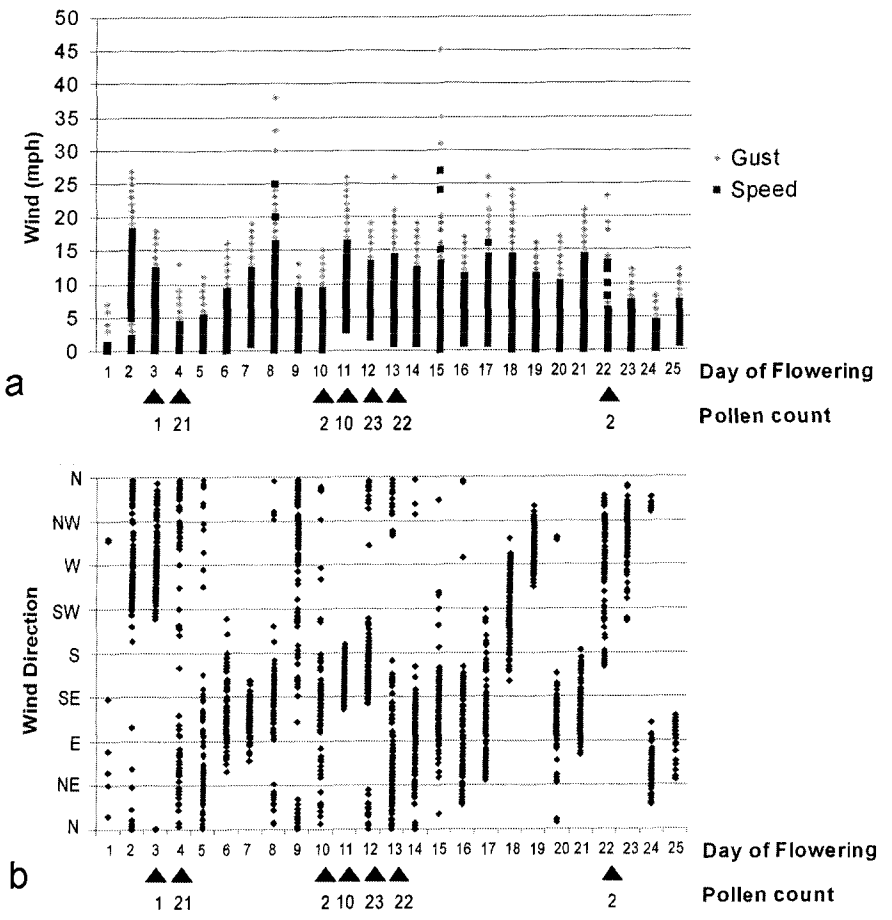


Figure 3. Graph of recorded wind speed (a) and wind direction (b) during flowering, the days when pollen was trapped '▲', and the number of pollen grains observed.

CHAPTER 6. GENERAL CONCLUSIONS

Molecular markers are now being routinely used in plant improvement programs and germplasm conservation activities. These markers, based on protein or DNA polymorphisms, provide additional information to observations on plant phenotypes. They are being utilized to help identify genes of interest, simplifying trait selection and diversity assessment. In the conservation of plant genetic resources, issues pertaining to germplasm management exist in many crops. As part of this dissertation, experiments were conducted to address curatorial concerns related to the *ex situ* conservation of oilseed *Brassica* such as the identification of life forms and verification of germplasm bulking strategies. The assessment of molecular variation, both on the genomic and single-locus scale, enabled us to gather information that may lead to the more effective management of *Brassica* germplasm at the NCRPIS and possibly help elucidate the control of flowering in rapeseed.

My first objective was to perform exploratory surveys of genetic variation by using microsatellite markers and DNA sequences to determine markers that could be associated to annual and biennial life forms of *B. napus*. The study was conceived in consultation with the crop curator to enable identification of life forms prior to planting and vernalization of *Brassica* accessions. My second objective was to validate specific genebank practices, such as germplasm bulking and the field regeneration of accessions in screen cages.

To achieve those objectives, I tested several hypotheses including 1) that *B. napus* life forms can be associated with microsatellite or DNA polymorphisms in the flowering-time gene, *FLC*, 2) that existing germplasm bulks of *B. napus* and *B. rapa* are composed of homogeneous component accessions, and 3) that contamination due to pollen flow is sufficiently precluded by the NCRPIS's current screen-cage regeneration protocol for *Brassica* to meet the current AOSCA seed purity standards.

Through testing of the first hypothesis (Chapter 2), microsatellite polymorphisms among a set of fifty representative accessions of *B. napus*, including both annual and biennial types, were identified. Microsatellite loci were selected for even representation of the 18 described *B. napus* linkage groups (Lowe et al., 2004). Molecular marker analyses were conducted in conjunction with the characterization of flowering time in the field without vernalization. Observations on flowering time in the absence of a vernalization treatment allowed the identification of true annual phenotypes of *B. napus*, complementing and improving the existing flowering data held in the Germplasm Resources Information Network database. Analysis of microsatellite marker profiles indicated that groups of life forms can be distinguished, as well as probable duplicate accessions. Results of the association analysis revealed eleven SSR loci with significant correlations to flowering time. Additional investigation is suggested to determine if these are proximal to or within genomic regions that control vernalization and flowering response, or simply in a weaker linkage disequilibrium with such regions. The analysis of molecular variation revealed that diversity 'within' life forms is greater than 'between' types of life forms in our selected set of representative accessions. It was also found out that there are statistically significant, but relatively weak, associations between derived genetic distances and geographic origins and between derived genetic distances and life forms.

In Chapter 3, I focused on two *FLOWERING LOCUS C* genes (*FLC1* and *FLC3*), which are central to flowering control and determined whether variation in nucleotide sequences at these loci exists between life forms. Using the same set of accessions as in the first study, I sequenced a region corresponding to exons 4 to 6 in all the representative accessions and to exons 2 to 7 in a selected subset of seven accessions. The shorter region analyzed corresponds to a gene region which is relatively conserved but known to be important in the functional properties of *FLC*. Our objective was to find polymorphisms in that region that could be associated with the need for vernalization. Results of the survey indicate that there were few polymorphisms and relatively little genetic differentiation between reproductive life forms, compared to the variation within life forms. When

case-control association tests were done, about thirty and six significant sites were determined in *FLC3* and *FLC1,* respectively. However, this information was inadequate to classify life forms. My observations of the *FLC* gene sequences are comparable to what was observed in published studies in *CONSTANS,* another highly conserved flowering-control gene in *Brassica*. My results based on data obtained directly from a gene responsible for flowering corroborates the low differentiation between the representative annual and biennial types detected with microsatellites in the first study.

The focus of the fourth and fifth chapters shift from exploratory surveys to examination of good genebanking practices. These studies provide information to challenge the assumptions behind current practices. The fourth chapter addresses issues in germplasm bulking in *B. rapa* and the problem of determining duplicate accessions in *B. napus*. Currently, these accessions are grouped based on morphological and phenological data. I examined if sufficient information can be derived from marker profiles from ten microsatellites. The molecular characterization data were analyzed with assignment tests utilizing a Bayesian framework, and with traditional methods, including tests for population differentiation and cluster analysis of genetic distances. The results of this study indicated that assignment tests together with the traditional methods can be used to verify or disprove the homogeneity of component accessions of bulks in *B. rapa*. It also addressed the problem of determination of duplicate accessions in *B. napus*. Unique accessions among *B. rapa* bulks and *B. napus* putative duplicates were identified. It was demonstrated that genetic-probability profiles of component accessions can corroborate cluster analyses, while providing additional information for decision making. Tests for population differentiation also added assurance to the uniqueness of component accessions in *B. rapa*. All of the tests provide the curator with information that assists in making well-informed management decisions on how to handle existing bulks and probable duplicates among the oilseed *Brassica* germplasm collection. Use of both molecular characterization and morphological and phenological information proved superior to use of any single type of information. For the assignment test method of analysis to be routinely used by the *Brassica* curator,

additional studies are recommended to establish baselines on acceptable assignment probability levels using subgroups of accessions with known levels of "breeding".

The fifth chapter reported results of tests to determine if pollen flow occurs between caged accessions of *B. napus* during field regeneration or is precluded, which impacts the ability of the program to maintain genetic integrity of the collections. In this chapter, molecular markers were not used, rather, a simple bioassay that detects expression of a dominant herbicide-resistance gene in the progeny of a non-herbicide resistant parent. This study also determined the size range of *B. napus* pollen and demonstrated that its pollen grains can pass through the screen mesh currently used by the NCRPIS and potentially cause contamination. However, the contamination rate between screened *B. napus* field plots during germplasm regeneration was very low, and within the current AOSCA limits for maintaining genetic identity. This supports the current method of regenerating *B. napus* in screen cages. However, I recommend follow-up tests to determine how pollen flow might vary in different years, if a finer screen or greater distance between cages could completely eliminate pollen flow, and if levels of cross-contamination vary across the more than 20 species of *Brassica* being conserved at the genebank.

The incorporation of molecular technologies into the germplasm management of *B. napus* and *B. rapa* are challenging and requires additional time for full time curators. However, the added information derived by using molecular characterization, as demonstrated in studies presented here, can be useful in decision-making. The cross-species amplification of most *Brassica* SSRs, together with the high variability that they generate, makes them suitable markers for molecular characterization that can support the classification of life forms, determination of identity and assessment of genetic integrity.

REFERENCES CITED

Adams, H., J.G. Vaughan, and G.R. Fenwick. 1989. The use of glucosinolates for cultivar identification in swede, *Brassica napus* L. var. *napobrassica* (L.) Peterm. Journal of the Science of Food and Agriculture 46:319-324.

Alvarez-Buylla, E., S. Pelaz, S.J. Liljegren, S.E. Gold, C. Burgedd, G.S. Ditta, L.R. de Pouplana, L. Martinez-Castilla, and M.F. Yanofsky. 2000. An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. Proc. Nat. Acad. Sci. (USA) 97:5328-5333.

Amasino, R.M. 2005. Vernalization and flowering time. Curr. Op. Biotech. 16:154-158.

Araki, T. 2001. Transition from vegetative to reproductive phase. Curr. Op. Plant Biol. 4:63-68.

Aswath, C.R. and S.H. Kim. 2005. Another story of MADS-box genes - their potential in plant biotechnology. Plant Growth Regulation 46:177-188.

Auld, D.L., K.A. Mahler, and A.A. Voorhis. 1988. Evaluation of the USDA collection of *Brassica* for fatty acid composition and glucosinolate content. Miscellaneous Series Bulletin No. 114. Agricultural Experiment Station, College of Agriculture, University of Idaho, Moscow, ID.

Avise, J.C. 2004. Molecular markers, natural history, and evolution. 2nd ed. Sinauer Associates, Inc., Sunderland, MA.

Axelsson, T, O Shavorskaya, and U Lagercrantz. 2001. Multiple flowering time QTLs within several *Brassica* species could be the result of duplicated copies of one ancestral gene. Genome 44:856-864

Barrier, M., R.H. Robichaux, and M.D. Purugganan. 2001. Accelerated regulatory gene evolution in an adaptive radiation. Proc. Natl. Acad. Sci. (USA) 98: 10208-10213.

Bastow, R., J.S. Myine, C. Lister, Z. Lippmann, R.A. Martienssen, and C. Dean. 2004. Vernalization requires epigenetic silencing of *FLC* by histone methylation. Nature 427:164-167.

Battey, N.H. 2000. Aspects of seasonality. J. Exp. Bot. 51:1769-1780.

Becker, H.C., C. Daamgard, and B. Karlsson. 1992. Environmental variation for outcrossing in rapeseed (*Brassica napus*). Theor. Appl. Genet. 84:303-306.

BioMat.Net. 2001. Biological Materials for Non Food products: FAIR Programme (FP4) Projects. Accessed online at http://www.nf-2000.org/secure/Fair/S335.htm and http://www.nf-2000.org/secure/Fair/F484.htm

Bretting, P.K. and M.P. Widrlechner. 1995. Genetic markers and plant genetic resource management. Pl. Br. Rev. 13: 11-86.

Butruille, D.V., R.P. Guries, and T.C. Osborn. 1999. Linkage analysis of molecular markers and quantitative trait loci in populations of inbred backcross lines of *Brassica napus* L. Genetics 153:949-964.

Casci, T. 2000. If winter comes, will flowers follow? Nature Rev. Genet. 1:166.

Charters, Y.M., A. Robertson, M.J. Wilkinson, and G. Ramsay. 1996. PCR analysis of oilseed rape cultivars (*Brassica napus* L. ssp. *oleifera*) using 5'-anchored simple sequence repeat (SSR) primers. Theor. Appl. Genet. 92:442-447.

Ciofi, C., S. M. Funk, T. Coote, D.J. Cheesman, R.L. Hammond, I.J. Saccheri, and M.W. Bruford. 1998. Genotyping with microsatellite markers. In: Molecular Tools for Screening Biodiversity. A. Karp, P. G. Isaac, and D.S. Ingram (eds.). Chapman & Hall, London, U.K. pp. 195-201.

Collard, B.C.Y., M.Z.Z. Jahufer, J.B. Brouwer, and E.C.K. Pang. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142:169–196.

Diaz, O., M. Gustafsson, and D. Astley. 1997. Effect of regeneration procedures on genetic diversity in *Brassica napus* and *B. rapa* as estimated by isozyme analysis. Genet. Res. Crop Evol. 44:523-532.

Downey, R.K. 1964. Selection of *Brassica campestris* L. containing no erucic acid in its seed oil. Can. J. Plant Sci. 44:295.

Doyle, M.R., S.J. Davis, R.M. Bastow, H.G. Mcwatters, L. Kozma-Bognár, F. Nagy, A.J. Millar, and R.M. Amasino. 2002. The *ELF4* gene controls circadian rhythms and flowering time in *Arabidopsis thaliana*. Nature 419:74-77.

Dubcovsky, J., A. Loukoianov, D. Fu, M. Valarik, A. Sanchez, and L. Yan. 2006. Effect of photoperiod on the regulation of wheat vernalization genes *VRN1* and *VRN2*. Plant Molecular Biology 60:469-480.

Eckardt, N.A. 2002. Alternative splicing and the control of flowering time. Plant Cell 14:743-747.

El-Assal, S.E., C. Alonso-Blanco, A.J.M. Peeters, V. Raz, and M. Koornneef. 2001. A QTL for flowering time in *Arabidopsis* reveals a novel allele for *CRY2*. Nature Genetics 29: 435-440.

Esquinas-Alcázar, J. 2005. Protecting crop genetic diversity for food security: Political, ethical and technical challenges. Nature Rev. Genetics 6:946-953.

Falk, K.C, G.F.W. Rakow, and R.K. Downey. 1998. The utilization of heterosis for seed yield in hybrid and synthetic cultivars of summer turnip rape. Can. J. Plant Sci. 78:383-387.

FAO. 2006. FAO agriculture data / FAOSTAT database. Food and Agriculture Organization of the United Nations. Accessed online at http://faostat.fao.org.

Farooq, S. and F. Azam. 2002. Molecular markers in plant breeding II. Some pre-requisites for use. Pakistan J. Biological Sci. 5:1141-1147.

Ferreira, M.E., J. Satagopan, B.S. Yandell, P.H. Williams, and T.C. Osborn. 1995. Mapping loci controlling vernalization requirement and flowering time in *Brassica napus*. Theor. Appl. Genet. 90:727-732.

Ford-Lloyd, B.V. 2001. Genotyping plant genetic resources. In: Plant Genotyping: The DNA Fingerprinting of Plants. R.J. Henry (ed.). CABI, Oxon, U.K. pp. 59-81.

Francia, E., G. Tacconi, C. Crosatti, D. Barabaschi, D. Bulgarelli, E. Dall'Aglio, and G. Valě. 2005. Marker assisted selection in crop plants. Plant Cell, Tissue and Organ Culture 82:317-342.

Gómez-Campo, C. 1999. Taxonomy. In: Biology of *Brassica* Coenospecies. (Developments in Plant Genetics and Breeding, vol. 4). C. Gómez-Campo (ed.). Elsevier Science, Amsterdam, The Netherlands. pp. 3-32.

Gómez-Campo, C. and S. Prakash. 1999. Origin and domestication. In: Biology of *Brassica* Coenospecies. (Developments in Plant Genetics and Breeding, vol. 4). C. Gómez-Campo (ed.). Elsevier Science, Amsterdam, The Netherlands. pp. 33-58.

Guil, J.L., I. Rodriguez-Garcia, and E. Torija. 1997. Nutritional and toxic factors in selected wild edible plants. Plant Foods for Human Nutrition 51: 99–107.

Gupta, V., A. Mukhopadhyay, N. Arumugam, Y.S. Sodhi, D. Pental, and A.K. Pradhan 2004. Molecular tagging of erucic acid trait in oilseed mustard (*Brassica juncea*) by QTL mapping and single nucleotide polymorphisms in *FAE1* gene. Theor. Appl. Genet. 108:743-749.

Hackenberg, E.M. and W. Kohler. 1996. Use of isozyme analysis in the breeding of synthetic rapeseed cultivars. Plant Breeding 115:474-479.

Hagenblad, J. and M. Nordborg. 2002. Sequence variation and haplotype structure surrounding the flowering time locus *FRI* in *Arabidopsis thaliana*. Genetics 161:289-298.

He, Y. and R. Amasino. 2004. Role of chromatin modification in flowering-time control. Trends in Plant Science 10:30-35.

Hecht, V, F. Foucher, C. Ferrandiz, R. Macknight, C. Navarro, J. Morin, M.E. Vardy, N. Ellis, J.P. Beltran, C. Rameau, and J.L. Weller. 2005. Conservation of *Arabidopsis* flowering genes in model legumes. Plant Physiology 137:1420-1434.

Henderson, I.R., F. Liu, S. Drea, G. Simpson, and C. Dean. 2005. An allelic series reveals essential roles for *FY* in plant development in addition to flowering-time control. Development 132: 3597-3607.

Holland, J.B., V.A. Portyanko, D.L. Hoffman, and M. Lee. 2002. Genomic regions controlling vernalization and photoperiod responses in oat. Theor. Appl. Genet. 105:113-126.

Hu, J., G. Li, D. Struss, and C.F. Quiros. 1999. SCAR and RAPD markers associated with 18-carbon fatty acids in rapeseed, *Brassica napus*. Plant Breeding 118:145.

IPGRI. 2001. Directory of germplasm collections: *Ex situ* germplasm holdings database. International Plant Genetic Resources Institute, Rome. Accessed online at http://ipgri.cgiar.org.

Johanson, U, J. West, C. Lister, S. Michaels, and R. Amasino, and C. Dean. 2000. Molecular analysis of *FRIGIDA*, a major determinant of natural variation in *Arabidopsis* flowering time. Science 290:344-347.

Kaufmann, K, R. Melzer, and G. Theißen. 2005. MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. Gene 347:183-198.

Kershaw, M.T. 1998. Ethno botanical leaflets: Canola. Accessed online at http://www.siu.edu/~ebl

Khlestkina, E.K. and E.A. Salina. 2006. SNP markers: Methods of analysis, ways of development, and comparison on an example of common wheat. Russian J. Genet. 42:585-594.

Kimber, D. and D.I. McGregor. 1995. *Brassica* Oilseeds: Production and Utilization. CABI, Wallingford, Oxfordshire, U.K.

Kole, C., P. Quijada, S.D. Michaels, R.M. Amasino, and T.C. Osborn. 2001. Evidence for homology of flowering-time genes *VFR2* from *Brassica rapa* and *FLC* from *Arabidopsis thaliana*. Theor. Appl. Genet. 102:425-430.

Koo, B., P.G. Pardey, and B. Wright. 2004. Saving Seeds: The Economics of Conserving Crop Genetic Resources *Ex Situ* in the Future Harvest Centres of the CGIAR. CABI, Wallingford, Oxfordshire, U.K.

Kruglyak, S., R.T. Durrett, M.D. Schug, and C.F. Aquadro. 1998. Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. Proc. Natl. Acad. Sci. (USA) 95:10774–10778.

Lackey, J. 1996. Biology of rapeseed. USDA – Animal and Plant Health Inspection Service (APHIS). Accessed online at http://www.aphis.usda.gov/biotech/rapeseed.html

Lagercrantz, U., M.K. Österberg, and M. Lascoux. 2002. Sequence variation and haplotype structure at the putative flowering-time locus *COL1* of *Brassica nigra* Mol. Biol. Evol. 19:1474-1482.

Lawton-Rauh, A.L., E.R. Alvarez-Buylla, and M.D. Purugganan. 2000. Molecular evolution of flower development. TREE 15:144-149.

Le Corre, V., F. Roux, and X. Reboud. 2002. DNA polymorphism at the *FRIGIDA* gene in *Arabidopsis thaliana*: Extensive nonsynonymous variation is consistent with local selection for flowering time. Mol. Biol. Evol. 19:1261-1271.

Li, Z., L. Zhao, C. Cui, G. Kai, L. Zhang, X. Sun, and K. Tang. 2005. Molecular cloning and characterization of an anti-bolting related gene (*BrpFLC*) from *Brassica rapa* ssp. *pekinensis.* Plant Science 168:407-413.

Lin, S., J. Wang, S. Poon, C. Su, S. Wang, and T. Chiou. 2005. Differential regulation of *FLOWERING LOCUS C* expression by vernalization in cabbage and *Arabidopsis*. Plant Physiology 137:1037-1048.

Love, C.G., J. Batley, G. Lim, A.J. Robinson, D. Savage, D. Singh, G.C. Spangenberg, and D. Edwards. 2004. New computational tools for *Brassica* genome research. Comp. Func. Genom. 5:276-280.

Lowe, A.J., A.E. Jones, A.F. Raybould, M. Trick, C.L. Moule, and K.J. Edwards. 2002. Transferability and genome specificity of new set of microsatellite primers among *Brassica* species of the U triangle. Molecular Ecology Notes 2:7-11.

Lowman, A.C. and M.D. Purugganan. 1999. Duplication of the *Brassica oleracea APETALA1* floral homeotic gene and the evolution of domesticated cauliflower. J. Hered. 90:514-520.

Macknight, R., M. Duroux, R. Laurie, P. Dijkwel, G. Simpson, and C. Dean. 2002. Functional significance of the alternative transcript processing of the *Arabidopsis* floral promoter *FCA*. Plant Cell 14:877-888.

Martynov, V.V. and E.E. Khavkin. 2004. Two homologs of the *FLOWERING LOCUS C* gene from leaf mustard (*Brassica juncea*). Russian J. Plant Physiol. 51:234-240.

Martynov, V.V. and E.E. Khavkin. 2005a. Polymorphism of the *CONSTANS* gene in *Brassica* plants. Russian J. Plant Physiol. 52:242-248.

Martynov, V.V. and E.E. Khavkin. 2005b. RFLP analysis of the *FLOWERING LOCUS C* gene in six *Brassica* species. Russian J. Plant Physiol. 52:352-258.

McNaughton, I.H. 1995. Swedes and rapes - *Brassica napus* (Cruciferae). In: Evolution of Crop Plants. 2nd ed. J. Smartt and N.W. Simmonds (eds.). Longman Scientific & Technical, Essex, U.K. pp. 68-75.

Meilan, R., A.M. Brunner, J.S. Skinner, and S.H. Strauss. 2001. Modification of flowering in transgenic trees. In: Molecular Breeding of Woody Plants: Proceedings of the International Wood Biotechnology Symposium (IWBS). Narita, Chiba, Japan, March 14-17, 2001. N. Moroshi and A. Komamine (eds.). Elsevier Science, Amsterdam, The Netherlands. pp. 247-256.

Michaels, S.D. and R.M. Amasino. 2000. Memories of winter: Vernalization and the competence to flower. Plant Cell & Environment 23:1145 -1154.

Michaels, S.D, I.C. Bezerra, and R.M. Amasino. 2004. *FRIGIDA*-related genes are required for the winter-annual habit in *Arabidopsis*. Proc. Natl. Acad. Sci. (USA) 101:3281-3285.

Michaels, S.D., E. Himelblau, S.Y. Kim, F.M. Schomburg, and R.M. Amasino. 2005. Integration of flowering signals in winter-annual *Arabidopsis*. Plant Physiology 136:149-156.

Mohammadi, S.A. and B.M. Prasanna. 2003. Analysis of genetic diversity in crop plants - Salient statistical tools and considerations. Crop Sci. 43:1235-1248.

Moon, J., H. Lee, M. Kim, and I. Lee. 2005. Analysis of flowering pathway integrators in *Arabidopsis*. Plant Cell Physiol. 46:292-299.

Mouradov, A., F. Crèmer, and G. Coupland. 2002. Control of flowering time: Interacting pathways as a basis for diversity. Plant Cell 14:S111-S130.

Mundges, H., W. Kohler, and W. Friedt. 1990. Identification of rape seed cultivars (*Brassica napus*) by starch-gel electrophoresis of enzymes. Euphytica 45:179-187.

Murphy, D.J. 2004. Genetic engineering of lipid composition. In: Biotechnology in Agriculture and Forestry, vol. 54: *Brassica.* E.C. Pua and C.J. Douglas (eds.). Springer-Verlag, Berlin, Germany. pp. 297-315.

Ng, M. and M.F. Yanofsky 2001. Function and evolution of the plant MADS-box gene family. Nature Rev. Genet. 2:186-195.

Olsen, K.M., A. Womack, A.R. Garrett, J.I. Suddith, and M.D. Purugganan. 2002. Contrasting evolutionary forces in the *Arabidopsis thaliana* floral development pathway. Genetics 160:1641-1650.

Osborn, T.C. 2004. The contribution of polyploidy to variation in *Brassica* species. Physiologia Plantarum 121:531-536.

Osborn, T.C. and L. Lukens. 2003. The molecular genetic basis of flowering time variation in *Brassica* species. In: Biotechnology in Agriculture and Forestry, Vol. 52: Brassicas and Legumes from Genome Structure to Breeding. T. Nagata and S. Tabata (eds.). Springer-Verlag, Berlin, Germany. pp. 69-86.

Osborn, T.C., C. Kole, I.A.P. Parkin, A.G. Sharpe, M. Kuiper, D.J. Lydiate, and M. Trick. 1997. Comparison of flowering time genes in *Brassica rapa, B. napus* and *Arabidopsis thaliana*. Genetics 146:1123-1129.

Österberg, M.K. O. Shavorskaya, M. Lascoux, and U. Lagercrantz. 2002. Naturally occurring indel variation in the *Brassica nigra COL1* gene is associated with variation in flowering time. Genetics 161:299-306.

Peterson, C.L., J.C. Thompson, J.S. Taberski, D.L. Reece, and G. Fleischman. 1999. Long-range on-road test with twenty-percent rapeseed biodiesel. Applied Engineering in Agriculture 15:91-101.

Phillips, P.W.B. and G.G. Khachatourians (eds.). 2001.The Biotechnology Revolution in Global Agriculture: Invention, Innovation and Investment in the Canola Sector. Biotechnology in Agriculture Series, no. 24. CABI, Wallingford, U.K.

Pires, J.C., J. Zhao, M.E. Schranz, E.J. Leon, P.A. Quijada, L.N. Lukens, and T.C. Osborn. 2004. Flowering time divergence and genomic rearrangements in resynthesized *Brassica* polyploids (Brassicaceae). Biol. J. Linn. Soc. 82:675-688.

Plant Biosafety Office. 1999. The biology of *Brassica rapa* L.: A companion document to the Directive 94-08, Assessment Criteria for Determining Environmental Safety of Plants with Novel Traits. Canadian Food Inspection Agency, Ottawa, Ontario. 20 pp.

Plieske, J. and D. Struss. 2001. Microsatellite markers for genome analysis in *Brassica*. I. development in *Brassica napus* and abundance in Brassicaceae species. Theor. Appl. Genet. 102:689-694.

Potts, D.A., G.W. Rakow, and D.R. Males. 1999. Canola quality *Brassica juncea*, a new oilseed crop for the Canadian Prairies. In: New Horizons for an Old Crop. Proceedings of the 10th International Rapeseed Congress. Canberra, Australia. Accessed online at http://www.regional.org.au/au/gcirc/4/70.htm#TopOfPage

Purugganan, M.D. 1997. The MADS-box floral homeotic gene lineages predate the origin of seed plants: Phylogenetic and molecular clock estimates. J. Mol. Evol. 45:392–396.

Purugganan, M.D., A.L. Boyles, and J.I. Suddith. 2000. Variation and selection at the *CAULIFLOWER* floral homeotic gene accompanying the evolution of domesticated *Brassica oleracea*. Genetics 155:855-862.

Purugganan, M.D., S.D. Rounsley, R.J. Schmidt, and M.F. Yanofsky. 1995. Molecular evolution of flower development: Diversification of the plant MADS-box regulatory gene family. Genetics 140: 345-356.

Putterill, J., R. Laurie, and R. MacKnight. 2004. It's time to flower: The genetic control of flowering time. BioEssays 26:363-373.

Quesada, V., C. Dean, and G.G. Simpson. 2005. Regulated RNA processing in the control of *Arabidopsis* flowering. Int. J. Dev. Biol. 49:773-780.

Rahman, M.H. 2001. Production of yellow-seeded *Brassica napus* through interspecific crosses. Plant Breeding 120:463-472.

Rahman, M.H. 2002. Fatty acid composition of resynthesized *Brassica napus* and trigenomic *Brassica* void of genes for erucic acid in their A genomes. Plant Breeding 121:357-359.

Ramanatha Rao, V. and T. Hodgkin. 2002. Genetic diversity and conservation and utilization of plant genetic resources. Plant Cell, Tissue, and Organ Culture 68:1-19.

Ratcliffe, O.J., R.W. Kumimoto, B.J. Wong, and J.L. Reichmann. 2003. Analysis of the *Arabidopsis MADS AFFECTING FLOWERING* gene family: *MAF2* prevents vernalization by short periods of cold. Plant Cell 15:1159-1169.

Ratcliffe, O.J., G.C. Nadzan, T.L. Reuber, and J.L. Riechmann. 2001. Regulation of flowering in *Arabidopsis* by an *FLC* homologue. Plant Physiol. 126:122-132.

Reichmann, J.L. and O.J. Ratcliffe. 2000. A genomic perspective on plant transcription factors. Curr. Op. Plant Biol. 3:423-434.

Rife, C. and H. Zeinali. 2003. Cold tolerance in oilseed rape over varying acclimation durations. Crop Sci. 43:96-100.

Rife, C.L. and J.P. Salgado. 1996. Selecting winter hardy oilseed rape for the great plains. In: Progress in New Crops. J. Janick (ed.). ASHS Press, Alexandria, VA. pp. 272-278.

Robert, L.S., F. Robson, A. Sharpe, D. Lydiate, and G. Coupland. 1998. Conserved structure and function of the *Arabidopsis* flowering time gene *CONSTANS* in *Brassica napus*. Plant Mol. Biol. 37:763-772.

Rudnóy, S., Z. Bratek, E. Páldi, I. Rácz, and D. Lásztity. 2002. *FLC*-like factors in wheat (cv. Mv15) and *Conyza* sp. Acta Biologica Szegediensis 46:45-46.

Sauer, J.D. 1993. Historical Geography of Crop Plants: A Select Roster. CRC Press, Boca Raton, FL.

Schlötterer, C. 2004. The evolution of molecular markers – just a matter of fashion? Nature Rev. Genetics 5:63-69.

Schrantz, M.E., P. Quijada, S. Sung, L. Lukens, R. Amasino, and T.C. Osborn. 2002. Characterization of the replicated flowering time gene *FLC* in *Brassica rapa*. Genetics 162:1457-1468.

Scortecci, K.C., S.D. Michaels, and R.M. Amasino. 2001. Identification of a MADS-box gene, *FLOWERING LOCUS M*, that represses flowering. Plant J. 26:229-236.

Shahidi, F. (ed.). 1990. Canola and Rapeseed: Production, Chemistry, Nutrition and Processing Technology. Van Nostrand Reinhold, New York, NY.

Sheldon, C.C., D.T. Rouse, E.J. Finnegan, W.J. Peacock, and E.S. Dennis. 2000. The molecular basis of vernalization: The central role of *FLOWERING LOCUS C* (*FLC*). Proc. Nat. Acad. Sci. (USA) 97:3753-3758.

Shindo, C., M.J. Aranzana, C. Lister, C. Baxter, C. Nicholls, M. Nordborg, and C. Dean. 2005. Role of *FRIGIDA* and *FLOWERING LOCUS C* in determining variation in flowering time of *Arabidopsis*. Plant Physiology 138:1163-1173.

Silva, C., J. Garcia-Mas, A.M. Sanchez, P. Arús, and M.M. Oliveira. 2005. Looking into flowering time in almond (*Prunus dulcis* (Mill.) D.A. Webb): The candidate gene approach. Theor. Appl. Genet. 110:959-968.

Simpson, G. 2005. NO flowering. BioEssays 27:239-241.

Simpson, G.G. and C. Dean. 2002. *Arabidopsis*, the Rosetta Stone of flowering time. Science 296: 285-289.

Simpson, G.G., A.R. Gendall, and C. Dean. 1999. When to switch to flowering. Annu. Rev. Cell Dev. Biol. 99:519-550.

Singh, B., J.N. Sachan, S.P. Singh, D.P. Pant, R.A. Khan, and R. Kumar. 2001. Correlation among fatty acids of *Brassica* and related species. Cruciferae Newsletter 23:9-10.

Spooner D., R. van Treuren, and M.C. de Vicente. 2005. Molecular markers for genebank management. IPGRI Technical Bulletin No. 10. International Plant Genetic Resources Institute, Rome, Italy.

Steffansson, B.R., F.W. Hougen, and R.K. Downey. 1961. Note on the isolation of rape plants with seed oil free from erucic acid. Can. J. Plant Sci. 41:218-219.

Stinchcombe, J.R., A.L. Caicedo, R. Hopkins, C. Mays, E.W. Boyd, M.D. Purugganan, and J. Schmitt. 2005. Vernalization in *Arabidopsis thaliana* (Brassicaceae): The effects of latitude and *FLC* variation. Am. J. Bot. 92:1701-1707.

Stinchcombe, J.R., C. Weinig, M. Ungerer, K.M. Olsen, C. Mays, S.S. Halldorsdottir, M.D. Purugganan, and J. Schmitt. 2004. A latitudinal cline in flowering time in *Arabidopsis thaliana* modulated by the flowering time gene *FRIGIDA*. Proc. Natl. Acad. Sci. (USA) 101:4712-4717.

Tadege, M., C.C. Sheldon, C.A. Helliwell, P. Stoutjesdijk, E.S. Dennis, and W.J. Peacock. 2001. Control of flowering time by *FLC* orthologues in *Brassica napus*. Plant J. 28:545-553.

Theissen, G., A. Becker, A. Di Rosa, A. Kanno, J.T. Kim, T. Munster, K. Winter, and H. Saedler. 2000. A short history of MADS-box genes in plants. Plant Mol. Biol. 42:115-149.

Thomas, P. 2003. Canola Grower's Manual. Canola Council of Canada,Winnipeg, Canada.

Tonguç, M. and P.D. Griffiths. 2004. Genetic relationships of *Brassica* vegetables determined using database derived simple sequence repeats. Euphytica 137:193-201.

Toxeopus, H., E. H. Oost, and G. Reuling. 1984. Current aspects of the taxonomy of cultivated *Brassica* species: The use of *B. rapa* L. versus *B. campestris* L. and a proposal for a new intraspecific classification of *B. rapa* L. Cruciferae Newsletter 9:55–57.

Tsunoda, S., K. Hinata, and C. Gómez-Campo (eds.). 1980. Brassica Crops and Wild Allies: Biology and Breeding. Japan Scientific Society Press, Tokyo.

Udall, J.A. 2003. A genetic study of oilseed *Brassica napus*: Mapping chromosome rearrangements and quantitative trait loci. Ph.D. dissertation. University of Wisconsin, Madison, WI.

Ungerer, M.C., S.S. Halldorsdottir, J.L. Modliszewski, T.F.C. Mackay, and M.D. Purugganan. 2002. Quantitative trait loci for inflorescence development in *Arabidopsis thaliana*. Genetics 160: 1133-1151.

USDA-GRIN. 2006. United States Department of Agriculture Germplasm Resources Information Network. Accessed online at http://www.ars-grin.gov/npgs

USDA-NASS. 2006. Crop Production 2005 Summary. National Agricultural Statistics Service. Accessed online at http://usda.mannlib.cornell.edu/reports/nassr/field/pcp-bban/cropan06.txt

Uzunova, M., W. Ecke, K. Weissleder, and G. Roebbelen. 1995. Mapping the genome of rapeseed (*Brassica napus* L.). 1. Construction of an RFLP linkage map and localization of QTLs for seed glucosinolate content. Theor. Appl. Genet. 90:194-204.

Vaugh, J.G. 1977. A multidisciplinary study of the taxonomy and origin of *Brassica* crops. BioScience 27:35-40.

Wang, R.L., A. Stec, J. Hey, L. Lukens, and J. Doebley. 1999. The limits of selection during maize domestication. Nature 398:236-239.

Warwick, S.I. and C.A. Sauder. 2005. Phylogeny of tribe Brassiceae (Brassicaceae) based on chloroplast restriction site polymorphisms and nuclear ribosomal internal transcribed spacer and chloroplast *trnL* intron sequences. Can. J. Bot. 83:467-483.

Werner, J.D., J.O. Borevitz, N.H. Uhlenhaut, J.R. Ecker, J. Chory, and D. Weigel. 2005. *FRIGIDA*-independent variation in flowering time of natural *Arabidopsis thaliana* accessions. Genetics 170:1197-1207.

White, J. and J.R. Law. 1991. Differentiation between varieties of oilseed rape (*Brassica napus* L.) on the basis of the fatty acid composition. Plant Var. Seeds 4:125-132.

Winichayakul, S., N.L. Beswick, C. Dean, and R.C. Macknight. 2005. Components of the *Arabidopsis* autonomous floral promotion pathway, *FCA* and *FY*, are conserved in monocots. Functional Plant Biology 32:345-355.

Yan, L., A. Loukoianov, A. Blechl, G. Tranquilli, W. Ramakrishna, P. SanMiguel, J.L. Bennetzen, V. Echenique, and J. Dubcovsky. 2004. The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. Science 303:1640–1644.

Yanovsky, M.J. and S.A. Kay. 2002. Molecular basis of seasonal time measurement in *Arabidopsis*. Nature 419:308-312.

Yanovsky, M.J. and S.A. Kay. 2003. Living by the calendar: How plants know when to flower. Nature Rev. Mol. Cell Biol. 4:265-275.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my major professors, Dr. Candice Gardner and Dr. E. Charles Brummer, and my program of study (POS) committee members, Dr. Kendall Lamkey, Dr. John Nason, and Dr. Mark Widrlechner for the guidance throughout my studies, and the helpful critiques of this dissertation. The molecular studies presented would not have been possible without the counsel of Dr. Randy Shoemaker and his generosity in allowing me to use his laboratory facilities. I'm indebted to Ms. Lori Lincoln, Ms. Jody Hayes, Dr. Diane Luth, and Dr. Jessica Schlueter for their valuable help in conducting the molecular marker research and DNA sequencing. I thank Mr. Larry Lockhart for facilitating the field and greenhouse research at the North Central Regional Plant Introduction Station (NCRPIS) and I highly appreciated the help of many NCRPIS staff and summer students during the course of the field and greenhouse experiments. For providing technical help and/or loaning to me field and laboratory equipment during my research, I thank Dr. Madan Bhattacharyya, Ms. Maria Hartt, Dr. David Grant, Ms. Corrinne Grover, Dr. Rex Nelson, Dr. Gary Polking, Dr. Devinder Sandhu, Mr. Mike Stahr, Dr. Shannon Schlueter, and Dr. Mark Westgate.

I would like to acknowledge the financial support from the United States Department of State, the Philippine Department of Agriculture, and the Iowa State Graduate College through the Fulbright student exchange program. I'm grateful to Ms. Marci Bushman, Ms. Jaci Severson, Ms. Susan Stolfus, Ms. Deb Vance, and Ms. Stacey Winter, for their help in coordinating my graduate program. For directly managing my scholarship grant, I thank the staff of the Philippine American Educational Foundation (PAEF) and the Institute of International Education (IIE) - Dr. Alexander Calata, Dr. Esmeralda Cunanan, Ms. Emily Guthridge, Ms. Mary Paschall, and Ms. Nora Yepes.

Finally, I'm extremely grateful to my family, Rocel, Maxine, and Mark, my parents, Lib and Tess, and sisters, Joy and Bing-Bing, for the constant encouragement, support, and prayers during the conduct of these studies and during our stay at Iowa State.